STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 148124

TO: Janet Epps-Ford **Location: REM/2C05/2C18** **Art Unit: 1635** **Friday, March 18, 2005** **Case Serial Number: 08/901612**	**From: David Schreiber** **Location: Biotech-Chem Library** **Remsen E01A61** **Phone: 272-2526** **david.schreiber@uspto.gov**

Search Notes

THIS PAGE LEFT BLANK

148124

Schreiber, David

From: Epps-Ford, Janet
Sent: Thursday, March 10, 2005 4:59 PM
To: Schreiber, David
Subject: sequence search request

PLease search SEQ ID NOS: 59-65 of application 08/901,612, each sequence is under 30 nucleotides in length. Search all pending and published nucleic acid sequence databases.

Thanks,
Janet L. Epps-Ford, Ph.D.
Art Unit 1635
Mailbox: Remsen 2C18
Office: Remsen 2C05
Phone: 571-272-0757
Fax: 571-273-0757

THIS PAGE LEFT BLANK

STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact ***the searcher or contact:***

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

- *I am an examiner in Workgroup:* ☐ *Example: 1610*
- *Relevant prior art **found**, search results used as follows:*
  - ☐ 102 rejection
  - ☐ 103 rejection
  - ☐ Cited as being of interest.
  - ☐ Helped examiner better understand the invention.
  - ☐ Helped examiner better understand the state of the art in their technology.

  Types of relevant prior art found:
  - ☐ Foreign Patent(s)
  - ☐ Non-Patent Literature
    (journal articles, conference proceedings, new product announcements etc.)
- *Relevant prior art **not found:***
  - ☐ Results verified the lack of relevant prior art (helped determine patentability).
  - ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library, Remsen Bldg.

THIS PAGE LEFT BLANK

Access DB# _____

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner # : _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: D. Schreiber
Searcher Phone #: 272-2526
Searcher Location: Remsen E01 A61
Date Searcher Picked Up: _____
Date Completed: 3/18
Searcher Prep & Review Time: 15
Clerical Prep Time: _____
Online Time 10

Type of Search

NA Sequence (#) 7
AA Sequence (#) _____
Structure (#) _____
Bibliographic _____
Litigation _____
Fulltext _____
Patent Family _____
Other _____

Vendors and cost where applicable

STN _____
Dialog _____
Questel/Orbit _____
Dr.Link _____
Lexis/Nexis _____
Sequence Systems Compugen
WWW/Internet _____
Other (specify) _____

PTO-1590 (8-01)

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 1025.6 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-59
Perfect score: 30
Sequence: 1 gacaugaacaagagaugauuaggcagaggt 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	30	100.0	30	6	AR027810	AR027810 Sequence
c 2	30	100.0	87	6	AX151115	AX151115 Sequence
c 3	30	100.0	99	14	HPBPRECA	M76687 Hepatitis B
c 4	30	100.0	99	14	HPBPRECB	M76688 Hepatitis B
c 5	30	100.0	99	14	HPBPRECC	M76689 Hepatitis B
c 6	30	100.0	99	14	HPBPRECD	M76690 Hepatitis B
c 7	30	100.0	99	14	HPBPRECE	M76691 Hepatitis B
c 8	30	100.0	99	14	HPBPRECF	M76692 Hepatitis B
c 9	30	100.0	99	14	HPBPRECG	M76693 Hepatitis B
c 10	30	100.0	99	14	HPBPRECH	M76694 Hepatitis B
c 11	30	100.0	99	14	HPBPRECI	M76695 Hepatitis B
c 12	30	100.0	99	14	HPBPRECM	M76699 Hepatitis B
c 13	30	100.0	129	6	AX151114	AX151114 Sequence
c 14	30	100.0	150	14	AF528205	AF528205 Hepatitis
c 15	30	100.0	150	14	AF528206	AF528206 Hepatitis
c 16	30	100.0	150	14	AF528207	AF528207 Hepatitis
c 17	30	100.0	150	14	AF528208	AF528208 Hepatitis
c 18	30	100.0	150	14	AF528209	AF528209 Hepatitis
c 19	30	100.0	150	14	AF528210	AF528210 Hepatitis
c 20	30	100.0	150	14	AF528211	AF528211 Hepatitis
c 21	30	100.0	150	14	AF528212	AF528212 Hepatitis
c 22	30	100.0	150	14	AF528213	AF528213 Hepatitis
c 23	30	100.0	150	14	AF528214	AF528214 Hepatitis
c 24	30	100.0	150	14	AF528215	AF528215 Hepatitis
c 25	30	100.0	150	14	AF528216	AF528216 Hepatitis
c 26	30	100.0	150	14	AF528217	AF528217 Hepatitis
c 27	30	100.0	150	14	AF528218	AF528218 Hepatitis
c 28	30	100.0	150	14	AF528219	AF528219 Hepatitis
c 29	30	100.0	150	14	AF528220	AF528220 Hepatitis
c 30	30	100.0	150	14	AF528221	AF528221 Hepatitis
c 31	30	100.0	150	14	AF528222	AF528222 Hepatitis
c 32	30	100.0	150	14	AF528224	AF528224 Hepatitis
c 33	30	100.0	150	14	AF528225	AF528225 Hepatitis
c 34	30	100.0	150	14	AF528226	AF528226 Hepatitis
c 35	30	100.0	150	14	AF528227	AF528227 Hepatitis
c 36	30	100.0	150	14	AF528228	AF528228 Hepatitis
c 37	30	100.0	150	14	AF528229	AF528229 Hepatitis
c 38	30	100.0	150	14	AF528231	AF528231 Hepatitis
c 39	30	100.0	150	14	AF528232	AF528232 Hepatitis
c 40	30	100.0	150	14	AF528233	AF528233 Hepatitis
c 41	30	100.0	150	14	AF528234	AF528234 Hepatitis
c 42	30	100.0	150	14	AF528235	AF528235 Hepatitis
c 43	30	100.0	150	14	AF528236	AF528236 Hepatitis
c 44	30	100.0	150	14	AF528237	AF528237 Hepatitis
c 45	30	100.0	150	14	AF528238	AF528238 Hepatitis

ALIGNMENTS

RESULT 1
AR027810
LOCUS AR027810 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 8 from patent US 5856459.
ACCESSION AR027810
VERSION AR027810.1 GI:5938630
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and Mills,J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 8 05-JAN-1999;
FEATURES Location/Qualifiers
source 1. .30
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 30;
Best Local Similarity 86.7%; Pred. No. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db          1 GACATGAACAAGAGATGATTAGGCAGAGGT 30
```

RESULT 2
AX151115/c
LOCUS AX151115 87 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 4 from Patent WO0138498.
ACCESSION AX151115
VERSION AX151115.1 GI:14533317
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F.,

```
          Fried,M. and Rossau,R.
  TITLE     A new genotype of hepatitis b virus
  JOURNAL   Patent: WO 0138498-A 4 31-MAY-2001;
          Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 6;  Length 87;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 3
HPBPRECA/c
LOCUS       HPBPRECA                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type1 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76687
VERSION     M76687.1  GI:485341
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45507.1"
                     /db_xref="GI:485342"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 1 (creates internal stop
                     codon)"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 4
HPBPRECB/c
LOCUS       HPBPRECB                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 2precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76688
VERSION     M76688.1  GI:485343
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     variation       2
                     /note="c in wt; t in virus type 2"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45508.1"
                     /db_xref="GI:485344"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 2 (creates internal stop
                     codon)"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 5
HPBPRECC/c
LOCUS       HPBPRECC                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 3precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76689
VERSION     M76689.1  GI:485345
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
```

```
                     /db_xref="taxon:10407"
     variation       6
                     /note="c in wt; t in virus type 3"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45509.1"
                     /db_xref="GI:485346"
                     /translation="MQLFHLCLIISCSCPTFQASKLCLGWL"
     variation       58
                     /gene="C"
                     /note="g in wt; t in virus type 3 (val to phe)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 3 (creates internal stop
                     codon)"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 6
HPBPRECD/c
LOCUS       HPBPRECD                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 4 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76690
VERSION     M76690.1  GI:485347
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45510.1"
                     /db_xref="GI:485348"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 4 (creates internal stop
                     codon)"
     variation       95
                     /note="g in wt; a in virus type 4 (gly to asp)"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 7
HPBPRECE/c
LOCUS       HPBPRECE                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 5 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76691
VERSION     M76691.1  GI:485349
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45511.1"
                     /db_xref="GI:485350"
                     /translation="MQLFHLCLIISCSCPTVQPSKLCLGWL"
     variation       64. .67
                     /gene="C"
                     /note="gcc in wt; ccg in virus type 5 (ala to pro)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 5 (creates internal stop
                     codon)"
     variation       95
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 8
HPBPRECF/c
LOCUS       HPBPRECF                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 6 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76692
VERSION     M76692.1  GI:485351
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
```

```
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .99
                     /gene="C"
     misc_feature    10. .99
                     /gene="C"
                     /product="precore protein"
                     /note="putative cds"
     variation       11
                     /gene="C"
                     /note="t in wt; c in virus type 6 (loss of start codon)"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 9
HPBPRECG/c
LOCUS       HPBPRECG                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 7 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76693
VERSION     M76693.1  GI:485352
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     misc_feature    10. .93
                     /gene="C"
                     /product="precore protein"
                     /standard_name="pre-C region note: putative CDS"
     variation       10
                     /gene="C"
                     /note="a in wt; t in virus type 7 (loss of start codon)"
     variation       14
                     /gene="C"
                     /note="a in wt; g in virus type 7 (gln to arg)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 7 ( creates internal stop
                     codon)"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 10
HPBPRECH/c
LOCUS       HPBPRECH                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 8 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76694
VERSION     M76694.1  GI:485353
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     misc_feature    10. .93
                     /gene="C"
                     /product="precore protein"
                     /standard_name="pre-C region note: putative CDS"
     variation       12
                     /gene="C"
                     /note="g in wt; t in virus type 8 (loss of start codon)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 8 (creates internal stop
                     codon)"
     variation       95
                     /note="g in wt; a in virus type 8 (gly to asp)"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 11
HPBPRECI/c
LOCUS       HPBPRECI                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 9 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76695
VERSION     M76695.1  GI:485354
```

```
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     misc_feature    10. .93
                     /gene="C"
                     /product="precore protein"
                     /standard_name="pre-C region note: putative CDS"
     variation       13
                     /gene="C"
                     /note="c in wt; t in virus type 9 (creates internal stop
                     codon)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 9 (creates internal stop
                     codon)"
     variation       95
                     /note="g in wt; a in virus type 9 (gly to asp)"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 12
HPBPRECM/c
LOCUS       HPBPRECM                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 13 precore protein (pre-C region, C) gene,
            5' end.
ACCESSION   M76699
VERSION     M76699.1  GI:485361
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .99
                     /gene="C"
     CDS             10. .>99
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45515.1"
                     /db_xref="GI:485362"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWLWDM"
     variation       95
                     /gene="C"
                     /note="g in wt; a in virus type 13 (gly to asp)"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 14;  Length 99;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||||:||::||||||||||
Db         52 GACATGAACAAGAGATGATTAGGCAGAGGT 23


RESULT 13
AX151114/c
LOCUS       AX151114                  129 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3 from Patent WO0138498.
ACCESSION   AX151114
VERSION     AX151114.1  GI:14533316
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F.,
            Fried,M. and Rossau,R.
  TITLE     A new genotype of hepatitis b virus
  JOURNAL   Patent: WO 0138498-A 3 31-MAY-2001;
            Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .129
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 6;  Length 129;
  Best Local Similarity   86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 14
AF528205/c
LOCUS       AF528205                  150 bp    DNA     linear   VRL 31-JUL-2003
DEFINITION  Hepatitis B virus ASC1123 core antigen precursor, gene, partial
            cds.
ACCESSION   AF528205
VERSION     AF528205.1  GI:32810971
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 150)
  AUTHORS   Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
  TITLE     Comparative evaluation of HBV precore and basal core promoter
            mutants in Indian patients with diverse clinical manifestations
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150)
  AUTHORS   Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUL-2002) Hepatitis Division, National Institute of
```

```
           Virology, 20-A, Dr Ambedkar Road, Pune, Maharashtra 411001, India
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="Hepatitis B virus"
                     /proviral
                     /mol_type="genomic DNA"
                     /isolate="ASC1123"
                     /isolation_source="asymptomatic HBsAg carrier"
                     /specific_host="Homo sapiens"
                     /db_xref="taxon:10407"
                     /country="India"
     misc_feature    <1. .>150
                     /note="contains partial basal core promoter"
     CDS             64. .>150
                     /note="contains complete precore region"
                     /codon_start=1
                     /product="core antigen precursor"
                     /protein_id="AAP87556.1"
                     /db_xref="GI:32810972"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWLXG"
ORIGIN

  Query Match           100.0%;  Score 30;  DB 14;  Length 150;
  Best Local Similarity  86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db        106 GACATGAACAAGAGATGATTAGGCAGAGGT 77


RESULT 15
AF528206/c
LOCUS       AF528206                150 bp    DNA     linear   VRL 31-JUL-2003
DEFINITION  Hepatitis B virus ASC1112 core antigen precursor, gene, partial
            cds.
ACCESSION   AF528206
VERSION     AF528206.1  GI:32810973
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 150)
  AUTHORS   Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
  TITLE     Comparative evaluation of HBV precore and basal core promoter
            mutants in Indian patients with diverse clinical manifestations
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150)
  AUTHORS   Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUL-2002) Hepatitis Division, National Institute of
            Virology, 20-A, Dr Ambedkar Road, Pune, Maharashtra 411001, India
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="Hepatitis B virus"
                     /proviral
                     /mol_type="genomic DNA"
                     /isolate="ASC1112"
                     /isolation_source="asymptomatic HBsAg carrier"
                     /specific_host="Homo sapiens"
                     /db_xref="taxon:10407"
                     /country="India"
     misc_feature    <1. .>150
                     /note="contains partial basal core promoter"
     CDS             64. .>150
                     /note="contains complete precore region"
                     /codon_start=1
                     /product="core antigen precursor"
                     /protein_id="AAP87557.1"
                     /db_xref="GI:32810974"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWLXG"
ORIGIN
```

```
  Query Match           100.0%;  Score 30;  DB 14;  Length 150;
  Best Local Similarity  86.7%;  Pred. No. 0.0036;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db        106 GACATGAACAAGAGATGATTAGGCAGAGGT 77


Search completed: March 17, 2005, 08:14:15
Job time : 1025.6 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 257 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-59
Perfect score: 30
Sequence: 1 gacaugaacaagagaugauuaggcagaggt 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30	100.0	30	2	AAT72562	Aat72562 Hepatitis
	2	30	100.0	30	2	AAT72563	Aat72563 Hepatitis
c	3	30	100.0	39	10	ADC64742	Adc64742 Hepatitis
c	4	30	100.0	87	4	AAD09094	Aad09094 Hepatitis
c	5	30	100.0	129	4	AAD09093	Aad09093 Hepatitis
c	6	30	100.0	639	6	AAD27422	Aad27422 Hepatitis
c	7	30	100.0	639	6	AAD31509	Aad31509 Hepatitis
c	8	30	100.0	655	4	AAH77569	Aah77569 HBV genot
c	9	30	100.0	655	4	AAH77568	Aah77568 HBV genot
c	10	30	100.0	655	4	AAH77574	Aah77574 HBV genot
c	11	30	100.0	655	4	AAH77573	Aah77573 HBV genot
c	12	30	100.0	655	4	AAH77570	Aah77570 HBV genot
c	13	30	100.0	655	4	AAH77571	Aah77571 HBV genot
c	14	30	100.0	664	4	AAH77572	Aah77572 HBV genot
c	15	30	100.0	669	12	ADO07220	Ado07220 Hepatitis
c	16	30	100.0	673	4	AAD09092	Aad09092 Hepatitis
c	17	30	100.0	675	4	AAH77563	Aah77563 HBV preCo
c	18	30	100.0	681	4	AAH77567	Aah77567 HBV genot
c	19	30	100.0	1395	2	AAV82688	Aav82688 Fulminant
c	20	30	100.0	1400	2	AAV82687	Aav82687 Fulminant
c	21	30	100.0	1445	2	AAV82692	Aav82692 Fulminant
c	22	30	100.0	1445	2	AAV82685	Aav82685 Fulminant
c	23	30	100.0	1445	2	AAV82690	Aav82690 Fulminant
c	24	30	100.0	1445	2	AAV82684	Aav82684 Fulminant
c	25	30	100.0	1500	2	AAV82695	Aav82695 Fulminant
c	26	30	100.0	1500	2	AAV82683	Aav82683 Fulminant
c	27	30	100.0	1500	2	AAV82694	Aav82694 Fulminant
c	28	30	100.0	1500	2	AAV82686	Aav82686 Fulminant
c	29	30	100.0	1500	2	AAV82706	Aav82706 Wild type
c	30	30	100.0	1500	2	AAV82689	Aav82689 Fulminant
c	31	30	100.0	1500	2	AAV82693	Aav82693 Fulminant
c	32	30	100.0	2342	1	AAN93072	Aan93072 Sequence
c	33	30	100.0	2743	1	AAN00003	Aan00003 Sequence
c	34	30	100.0	2743	2	AAQ04799	Aaq04799 Recombina
c	35	30	100.0	3180	4	AAH42375	Aah42375 Nucleotid
c	36	30	100.0	3182	6	AAD31765	Aad31765 Hepatitis
c	37	30	100.0	3182	9	ACA62422	Aca62422 Hepatitis
c	38	30	100.0	3182	10	AAD60866	Aad60866 Hepatitis
c	39	30	100.0	3220	3	AAZ88924	Aaz88924 Hepatitis
c	40	30	100.0	3248	4	AAD09091	Aad09091 Hepatitis
c	41	30	100.0	3248	4	AAH77562	Aah77562 HBV genot
c	42	30	100.0	5618	2	AAQ88310	Aaq88310 Plasmid p
c	43	30	100.0	7991	6	AAS16094	Aas16094 HBV viral
c	44	30	100.0	8007	6	AAS16092	Aas16092 HBV viral
c	45	30	100.0	8717	6	AAS16093	Aas16093 HBV viral

ALIGNMENTS

RESULT 1
AAT72562
ID AAT72562 standard; DNA; 30 BP.
XX
AC AAT72562;
XX
DT 03-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV88b.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1. .30
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
XX
PN WO9639502-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP002432.
XX
PR 06-JUN-1995; 95US-00467397.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (HYBR-) HYBRIDON INC.
XX
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX
DR WPI; 1997-043124/04.
XX
PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
PS Claim 1; Page 12; 81pp; English.
XX
CC The present sequence represents a synthetic oligonucleotide HBV88b which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a

```
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 10 G; 5 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 30;  DB 2;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 0.003;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||:||::||||||||||
Db          1 GACATGAACAAGAGATGATTAGGCAGAGGT 30


RESULT 2
AAT72563
ID   AAT72563 standard; DNA; 30 BP.
XX
AC   AAT72563;
XX
DT   03-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV88Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        1. .20
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   1
FT                   /*tag=  c
FT                   /mod_base= gm
FT   modified_base   2
FT                   /*tag=  d
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   3
FT                   /*tag=  e
FT                   /mod_base= cm
FT   modified_base   4
FT                   /*tag=  f
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   5
FT                   /*tag=  g
FT                   /mod_base= um
FT   modified_base   6
FT                   /*tag=  h
FT                   /mod_base= gm
FT   modified_base   7
FT                   /*tag=  i
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   8
FT                   /*tag=  j
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   9
FT                   /*tag=  k
FT                   /mod_base= cm
FT   modified_base   10
FT                   /*tag=  l
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   11
FT                   /*tag=  m
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   12
FT                   /*tag=  n
FT                   /mod_base= gm
FT   modified_base   13
FT                   /*tag=  o
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   14
FT                   /*tag=  p
FT                   /mod_base= gm
FT   modified_base   15
FT                   /*tag=  q
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   16
FT                   /*tag=  r
FT                   /mod_base= um
FT   modified_base   17
FT                   /*tag=  s
FT                   /mod_base= gm
FT   modified_base   18
FT                   /*tag=  s
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   19
FT                   /*tag=  u
FT                   /mod_base= um
FT   modified_base   20
FT                   /*tag=  v
FT                   /mod_base= um
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV88Mb which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 10 G; 1 T; 4 U; 0 Other;

  Query Match             100.0%;  Score 30;  DB 2;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 0.003;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||||||||||||||||||||||||||||
Db          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
```

```
RESULT 3
ADC64742/c
ID   ADC64742 standard; RNA; 39 BP.
XX
AC   ADC64742;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Hepatitis B virus DNA polymerase related RNA oligonucleotide.
XX
KW   screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
XX
OS   Synthetic.
OS   Hepatitis B virus.
XX
PN   KR2002007891-A.
XX
PD   29-JAN-2002.
XX
PF   19-JUL-2000; 2000KR-00041420.
XX
PR   19-JUL-2000; 2000KR-00041420.
XX
PA   (MOGA-) MOGAM BIOTECHNOLOGY INST.
PA   (VIRO-) VIROGEN CO LTD.
XX
PI   Ji HJ,  Jung SI,  Kim YC,  Min MG,  Ryu WS,  Yoon GS;
XX
DR   WPI; 2003-309015/30.
XX
PT   Screening of antiviral agents by protein-priming activity of hepatitis B
PT   virus DNA polymerase.
XX
PS   Disclosure; Page 12; 13pp; Korean.
XX
CC   The present invention describes a method of screening for an antiviral
CC   agent by the protein-priming activity of hepatitis B virus (HBV) DNA
CC   polymerase. Also described is developing an antiviral agent with a high
CC   selectivity to HBV which can be used for high-throughput screening. The
CC   present sequence represents an RNA oligonucleotide which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 39 BP; 5 A; 13 C; 3 G; 0 T; 18 U; 0 Other;

  Query Match            100.0%;  Score 30;  DB 10;  Length 39;
  Best Local Similarity   86.7%;  Pred. No. 0.0031;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||||:||::||||||||||
Db         37 GACATGAACAAGAGATGATTAGGCAGAGGT 8


RESULT 4
AAD09094/c
ID   AAD09094 standard; DNA; 87 BP.
XX
AC   AAD09094;
XX
DT   04-SEP-2001  (first entry)
XX
DE   Hepatitis B virus FR1 strain genotype G HBeAg DNA fragment.
XX
KW   HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
KW   preS2; surface antigen; HBsAg; HBX protein; vaccine; HBeAg;
KW   liver disease; hepatitis; liver cancer; HBcAg; core antigen; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200138498-A2.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US032108.
XX
PR   24-NOV-1999;   99US-0167206P.
XX
PA   (PHAR-) PHARMASSET INC.
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Schinazi R,  De Gendt S,  Van Geyt C,  Zoulim F,  Fried M;
PI   Rossau R;
XX
DR   WPI; 2001-367676/38.
XX
PT   Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT   polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT   treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
PS   Claim 6; Page 57; 84pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) strain FR1,
CC   genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
CC   PreS2 and surface antigen HBsAg) and HBX proteins. HBV genotype G nucleic
CC   acids and polypeptides are useful for diagnosing, prognosing and treating
CC   infections caused by HBV genotype G. They can be used in a vaccine to
CC   treat or prevent HBV genotype G infection. The HBV genotype G derived
CC   nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC   sample or diagnosis of HBV genotype G infection. The presence of HBV
CC   genotype G statistically correlates with the presence of liver damage
CC   and/or liver cancer in the subject. The HBV genotype G core insert
CC   peptide encoding nucleic acid is useful for designing monitoring assays
CC   to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC   and HBeAg (genotype G e antigen) in patients infected with HBV. The
CC   antibodies or antigens of HBV genotype G are useful for identifying a
CC   stage of liver disease caused by HBV genotype G. The present sequence is
CC   a hepatitis B virus (HBV) strain FR1, genotype G DNA fragment encoding e
CC   antigen (HBeAg)
XX
SQ   Sequence 87 BP; 14 A; 24 C; 17 G; 32 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 30;  DB 4;  Length 87;
  Best Local Similarity   86.7%;  Pred. No. 0.0035;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 5
AAD09093/c
ID   AAD09093 standard; DNA; 129 BP.
XX
AC   AAD09093;
XX
DT   04-SEP-2001  (first entry)
XX
DE   Hepatitis B virus FR1 strain genotype G DNA fragment #1.
XX
KW   HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
KW   preS2; surface antigen; HBsAg; HBX protein; vaccine; liver disease;
KW   hepatitis; liver cancer; HBcAg; core antigen; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200138498-A2.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US032108.
XX
PR   24-NOV-1999;   99US-0167206P.
XX
PA   (PHAR-) PHARMASSET INC.
```

```
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Schinazi R,  De Gendt S,  Van Geyt C,  Zoulim F,  Fried M;
PI   Rossau R;
XX
DR   WPI; 2001-367676/38.
XX
PT   Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT   polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT   treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
PS   Claim 5; Page 57; 84pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) strain FR1,
CC   genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
CC   PreS2 and surface antigen HBsAg) and HBX proteins. HBV genotype G nucleic
CC   acids and polypeptides are useful for diagnosing, prognosing and treating
CC   infections caused by HBV genotype G. They can be used in a vaccine to
CC   treat or prevent HBV genotype G infection. The HBV genotype G derived
CC   nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC   sample or diagnosis of HBV genotype G infection. The presence of HBV
CC   genotype G statistically correlates with the presence of liver damage
CC   and/or liver cancer in the subject. The HBV genotype G core insert
CC   peptide encoding nucleic acid is useful for designing monitoring assays
CC   to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC   and HBeAg (genotype G e antigen) in patients infected with HBV. The
CC   antibodies or antigens of HBV genotype G are useful for identifying a
CC   stage of liver disease caused by HBV genotype G. The present sequence is
CC   a hepatitis B virus (HBV) strain FR1, genotype G DNA fragment
XX
SQ   Sequence 129 BP; 25 A; 32 C; 26 G; 46 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 30;  DB 4;  Length 129;
  Best Local Similarity   86.7%;  Pred. No. 0.0038;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 6
AAD27422/c
ID   AAD27422 standard; DNA; 639 BP.
XX
AC   AAD27422;
XX
DT   18-APR-2002  (first entry)
XX
DE   Hepatitis B virus (HBV) core antigen (HBcAg) encoding DNA #1.
XX
KW   Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
KW   prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KW   hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW   tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KW   dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KW   food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
KW   antiprotozoal; ds.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag=  a
FT                   /product= "HBcAg"
XX
PN   WO200198333-A2.
XX
PD   27-DEC-2001.
XX
PF   22-JUN-2001; 2001WO-GB002817.
XX
PR   22-JUN-2000; 2000GB-00015308.
PR   06-OCT-2000; 2000GB-00024544.
XX
PA   (CELL-) CELLTECH PHARM LTD.
XX
PI   Page M,  Li J,  Pumpens P;
XX
DR   WPI; 2002-098223/13.
DR   P-PSDB; AAE17018.
XX
PT   New proteins comprising a modified hepatitis B core antigen, useful as a
PT   vaccine in prophylactic or therapeutic vaccination of the human or animal
PT   body, particularly against hepatitis B virus infection.
XX
PS   Disclosure; Page 38-39; 40pp; English.
XX
CC   The invention relates to modified proteins comprising hepatitis B virus
CC   (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC   repeats has been deleted and the protein comprising the C-terminal
CC   cysteine of HBcAg. The deleted region may be replaced by an epitope from
CC   a protein other than HBcAg, in which case the HBcAg acts as a carrier to
CC   present the epitope to the immune system. This chimeric protein or its
CC   nucleic acid is useful as a vaccine or in a method of prophylactic or
CC   therapeutic vaccination of the human or animal body, particularly against
CC   HBV. The nucleic acid encoding the protein may be used in gene therapy or
CC   DNA vaccination protocols. The chimeric protein or its nucleic acid may
CC   also be used as the basis of a prophylactic vaccine against a range of
CC   diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
CC   influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
CC   immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
CC   tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
CC   diarrhoea, meningitis or gonorrhea. The present sequence is a DNA
CC   encoding Hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 30;  DB 6;  Length 639;
  Best Local Similarity   86.7%;  Pred. No. 0.0049;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 7
AAD31509/c
ID   AAD31509 standard; DNA; 639 BP.
XX
AC   AAD31509;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Hepatitis B virus core antigen (HBcAg) encoding DNA.
XX
KW   Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
KW   therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
KW   polio; herpes; rabies; AIDS; foot-and-mouth disease; ds.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag=  a
FT                   /product= "HBc protein"
FT   sig_peptide     1. .87
FT                   /*tag=  b
FT   mat_peptide     88. .636
FT                   /*tag=  c
FT                   /product= "Mature HBc protein"
XX
PN   WO200177158-A1.
XX
PD   18-OCT-2001.
```

```
XX
PF   09-APR-2001; 2001WO-GB001607.
XX
PR   07-APR-2000; 2000EP-00107118.
XX
PA   (MEDE-) MEDEVA EURO LTD.
XX
PI   Gehin A,  Gilbert R,  Stuart D,  Rowlands D;
XX
DR   WPI; 2002-239995/29.
DR   P-PSDB; AAE19793.
XX
PT   Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for the
PT   prophylactic or therapeutic treatment of humans or animals against e.g.
PT   HB virus, viral hepatitis, hepatitis C virus, influenza, or foot-and-
PT   mouth disease.
XX
PS   Disclosure; Page 23-24; 27pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) core antigen
CC   (HBcAg) fusion proteins and polynucleotides encoding such proteins.
CC   Sequences of the invention are useful in methods of prophylactic or
CC   therapeutic vaccination or to manufacture medicaments for prophylactic or
CC   therapeutic vaccination of the human or animal body against HBV, e.g.
CC   against viral hepatitis. They are also useful as a prophylactic vaccine
CC   against e.g. hepatitis c virus, influenza, polio, herpes, rabies,
CC   acquired immune deficiency syndrome (AIDS) or foot-and-mouth disease. The
CC   present sequence is a DNA encoding hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 30;  DB 6;  Length 639;
  Best Local Similarity   86.7%;  Pred. No. 0.0049;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 8
AAH77569/c
ID   AAH77569 standard; DNA; 655 BP.
XX
AC   AAH77569;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV genotype G strain US1 preCore/Core DNA.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW   HBeAg; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP011526.
XX
PR   03-DEC-1999;   99EP-00870252.
PR   07-DEC-1999;   99US-0169287P.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Van Geyt C,  De Gendt S;
XX
DR   WPI; 2001-374785/39.
XX
PT   Novel isolated and/or purified hepatitis B virus polypeptide and
PT   polynucleotide sequences that are phylogenetically different from HBV
PT   genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT   therapy.
XX
PS   Claim 3; Fig 7; 94pp; English.
XX
CC   The invention relates to the complete nucleic acid sequence of a new
CC   human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC   This genotype was found with a high prevalence in patients chronically
CC   infected with HBV and residing in Europe and the USA. The invention
CC   relates to a fully defined sequence of 3248 nucleotides as given in
CC   specification, a sequence with 92% identity to the given sequence, or
CC   sequence that is degenerate to the mentioned sequences. These
CC   polynucleotides are useful for HBV genotyping. The proteins encoded by
CC   the polynucleotides are useful for detecting antibodies in a biological
CC   sample. Ligands that bind to the proteins and antibodies directed against
CC   the proteins are useful for detecting the proteins and for detecting
CC   HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC   preparing a vaccine or medicament for treating HBV infections. The
CC   present sequence is provided in an alignment of preCore/Core sequences of
CC   an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC   US6, US7, US9, US10) of HBV genotype G
XX
SQ   Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

  Query Match             100.0%;  Score 30;  DB 4;  Length 655;
  Best Local Similarity   86.7%;  Pred. No. 0.005;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 9
AAH77568/c
ID   AAH77568 standard; DNA; 655 BP.
XX
AC   AAH77568;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV genotype G strain FR2 preCore/Core DNA.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW   HBeAg; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP011526.
XX
PR   03-DEC-1999;   99EP-00870252.
PR   07-DEC-1999;   99US-0169287P.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Van Geyt C,  De Gendt S;
XX
DR   WPI; 2001-374785/39.
XX
PT   Novel isolated and/or purified hepatitis B virus polypeptide and
PT   polynucleotide sequences that are phylogenetically different from HBV
PT   genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT   therapy.
XX
PS   Claim 3; Fig 7; 94pp; English.
XX
CC   The invention relates to the complete nucleic acid sequence of a new
CC   human hepatitis B virus (HBV) genotype, provisionally named genotype G.
```

CC This genotype was found with a high prevalence in patients chronically
CC infected with HBV and residing in Europe and the USA. The invention
CC relates to a fully defined sequence of 3248 nucleotides as given in
CC specification, a sequence with 92% identity to the given sequence, or
CC sequence that is degenerate to the mentioned sequences. These
CC polynucleotides are useful for HBV genotyping. The proteins encoded by
CC the polynucleotides are useful for detecting antibodies in a biological
CC sample. Ligands that bind to the proteins and antibodies directed against
CC the proteins are useful for detecting the proteins and for detecting
CC HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC preparing a vaccine or medicament for treating HBV infections. The
CC present sequence is provided in an alignment of preCore/Core sequences of
CC an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC US6, US7, US9, US10) of HBV genotype G
XX
SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 30; DB 4; Length 655;
Best Local Similarity 86.7%; Pred. No. 0.005;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14
```

RESULT 10
AAH77574/c
ID AAH77574 standard; DNA; 655 BP.
XX
AC AAH77574;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype G strain US10 preCore/Core DNA.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW HBeAg; ds.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP011526.
XX
PR 03-DEC-1999; 99EP-00870252.
PR 07-DEC-1999; 99US-0169287P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.
XX
PS Claim 3; Fig 7; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC This genotype was found with a high prevalence in patients chronically
CC infected with HBV and residing in Europe and the USA. The invention
CC relates to a fully defined sequence of 3248 nucleotides as given in
CC specification, a sequence with 92% identity to the given sequence, or
CC sequence that is degenerate to the mentioned sequences. These
CC polynucleotides are useful for HBV genotyping. The proteins encoded by
CC the polynucleotides are useful for detecting antibodies in a biological
CC sample. Ligands that bind to the proteins and antibodies directed against
CC the proteins are useful for detecting the proteins and for detecting
CC HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC preparing a vaccine or medicament for treating HBV infections. The
CC present sequence is provided in an alignment of preCore/Core sequences of
CC an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC US6, US7, US9, US10) of HBV genotype G
XX
SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 30; DB 4; Length 655;
Best Local Similarity 86.7%; Pred. No. 0.005;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14
```

RESULT 11
AAH77573/c
ID AAH77573 standard; DNA; 655 BP.
XX
AC AAH77573;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype G strain US7 preCore/Core DNA.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW HBeAg; ds.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP011526.
XX
PR 03-DEC-1999; 99EP-00870252.
PR 07-DEC-1999; 99US-0169287P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.
XX
PS Claim 3; Fig 7; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC This genotype was found with a high prevalence in patients chronically
CC infected with HBV and residing in Europe and the USA. The invention
CC relates to a fully defined sequence of 3248 nucleotides as given in
CC specification, a sequence with 92% identity to the given sequence, or
CC sequence that is degenerate to the mentioned sequences. These
CC polynucleotides are useful for HBV genotyping. The proteins encoded by
CC the polynucleotides are useful for detecting antibodies in a biological
CC sample. Ligands that bind to the proteins and antibodies directed against
CC the proteins are useful for detecting the proteins and for detecting
CC HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC preparing a vaccine or medicament for treating HBV infections. The
CC present sequence is provided in an alignment of preCore/Core sequences of
CC an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC US6, US7, US9, US10) of HBV genotype G

```
XX
SQ   Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

  Query Match            100.0%;  Score 30;  DB 4;  Length 655;
  Best Local Similarity   86.7%;  Pred. No. 0.005;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 12
AAH77570/c
ID   AAH77570 standard; DNA; 655 BP.
XX
AC   AAH77570;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV genotype G strain US3 preCore/Core DNA.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW   HBeAg; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP011526.
XX
PR   03-DEC-1999;   99EP-00870252.
PR   07-DEC-1999;   99US-0169287P.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Van Geyt C,  De Gendt S;
XX
DR   WPI; 2001-374785/39.
XX
PT   Novel isolated and/or purified hepatitis B virus polypeptide and
PT   polynucleotide sequences that are phylogenetically different from HBV
PT   genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT   therapy.
XX
PS   Claim 3; Fig 7; 94pp; English.
XX
CC   The invention relates to the complete nucleic acid sequence of a new
CC   human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC   This genotype was found with a high prevalence in patients chronically
CC   infected with HBV and residing in Europe and the USA. The invention
CC   relates to a fully defined sequence of 3248 nucleotides as given in
CC   specification, a sequence with 92% identity to the given sequence, or
CC   sequence that is degenerate to the mentioned sequences. These
CC   polynucleotides are useful for HBV genotyping. The proteins encoded by
CC   the polynucleotides are useful for detecting antibodies in a biological
CC   sample. Ligands that bind to the proteins and antibodies directed against
CC   the proteins are useful for detecting the proteins and for detecting
CC   HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC   preparing a vaccine or medicament for treating HBV infections. The
CC   present sequence is provided in an alignment of preCore/Core sequences of
CC   an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC   US6, US7, US9, US10) of HBV genotype G
XX
SQ   Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

  Query Match            100.0%;  Score 30;  DB 4;  Length 655;
  Best Local Similarity   86.7%;  Pred. No. 0.005;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 13
AAH77571/c
ID   AAH77571 standard; DNA; 655 BP.
XX
AC   AAH77571;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV genotype G strain US5 preCore/Core DNA.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW   HBeAg; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP011526.
XX
PR   03-DEC-1999;   99EP-00870252.
PR   07-DEC-1999;   99US-0169287P.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Van Geyt C,  De Gendt S;
XX
DR   WPI; 2001-374785/39.
XX
PT   Novel isolated and/or purified hepatitis B virus polypeptide and
PT   polynucleotide sequences that are phylogenetically different from HBV
PT   genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT   therapy.
XX
PS   Claim 3; Fig 7; 94pp; English.
XX
CC   The invention relates to the complete nucleic acid sequence of a new
CC   human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC   This genotype was found with a high prevalence in patients chronically
CC   infected with HBV and residing in Europe and the USA. The invention
CC   relates to a fully defined sequence of 3248 nucleotides as given in
CC   specification, a sequence with 92% identity to the given sequence, or
CC   sequence that is degenerate to the mentioned sequences. These
CC   polynucleotides are useful for HBV genotyping. The proteins encoded by
CC   the polynucleotides are useful for detecting antibodies in a biological
CC   sample. Ligands that bind to the proteins and antibodies directed against
CC   the proteins are useful for detecting the proteins and for detecting
CC   HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC   preparing a vaccine or medicament for treating HBV infections. The
CC   present sequence is provided in an alignment of preCore/Core sequences of
CC   an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC   US6, US7, US9, US10) of HBV genotype G
XX
SQ   Sequence 655 BP; 138 A; 154 C; 140 G; 195 T; 0 U; 28 Other;

  Query Match            100.0%;  Score 30;  DB 4;  Length 655;
  Best Local Similarity   86.7%;  Pred. No. 0.005;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 14
AAH77572/c
```

```
ID   AAH77572 standard; DNA; 664 BP.
XX
AC   AAH77572;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV genotype G strain US6 preCore/Core DNA.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW   HBeAg; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP011526.
XX
PR   03-DEC-1999;   99EP-00870252.
PR   07-DEC-1999;   99US-0169287P.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Van Geyt C,  De Gendt S;
XX
DR   WPI; 2001-374785/39.
XX
PT   Novel isolated and/or purified hepatitis B virus polypeptide and
PT   polynucleotide sequences that are phylogenetically different from HBV
PT   genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT   therapy.
XX
PS   Claim 3; Fig 7; 94pp; English.
XX
CC   The invention relates to the complete nucleic acid sequence of a new
CC   human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC   This genotype was found with a high prevalence in patients chronically
CC   infected with HBV and residing in Europe and the USA. The invention
CC   relates to a fully defined sequence of 3248 nucleotides as given in
CC   specification, a sequence with 92% identity to the given sequence, or
CC   sequence that is degenerate to the mentioned sequences. These
CC   polynucleotides are useful for HBV genotyping. The proteins encoded by
CC   the polynucleotides are useful for detecting antibodies in a biological
CC   sample. Ligands that bind to the proteins and antibodies directed against
CC   the proteins are useful for detecting the proteins and for detecting
CC   HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC   preparing a vaccine or medicament for treating HBV infections. The
CC   present sequence is provided in an alignment of preCore/Core sequences of
CC   an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC   US6, US7, US9, US10) of HBV genotype G
XX
SQ   Sequence 664 BP; 146 A; 160 C; 144 G; 208 T; 0 U; 6 Other;

  Query Match             100.0%;  Score 30;  DB 4;  Length 664;
  Best Local Similarity   86.7%;  Pred. No. 0.005;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||||:||::||||||||||
Db         43 GACATGAACAAGAGATGATTAGGCAGAGGT 14


RESULT 15
ADO07220/c
ID   ADO07220 standard; DNA; 669 BP.
XX
AC   ADO07220;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Hepatitis B virus core antigen DNA.
XX
KW   HBcAg; immunomodulator; vaccine; gene; ss.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             10. .669
FT                   /*tag=  a
FT                   /product= "HBcAg"
FT                   /partial
FT                   /note= "No start codon"
XX
PN   WO2004035007-A2.
XX
PD   29-APR-2004.
XX
PF   17-OCT-2003; 2003WO-US033178.
XX
PR   17-OCT-2002; 2002US-0419279P.
XX
PA   (ORAG-) ORAGEN CORP.
XX
PI   Michaels F;
XX
DR   WPI; 2004-348329/32.
DR   P-PSDB; ADO07221.
XX
PT   Modulating a systemic immune response to a peptide in a mammal comprises
PT   transmucosally administering a macromolecular aggregate of the peptide.
XX
PS   Disclosure; SEQ ID NO 1; 81pp; English.
XX
CC   The present sequence is the DNA sequence of the hepatitis B virus core
CC   antigen (HBcAg) gene from HBV serotype ayw. A peptide comprising a HBV
CC   protein can be used in claimed methods of the invention for modulating an
CC   immune response in a mammal. A method of inducing a systemic immune
CC   response to a peptide in a mammal comprises transmucosally administering
CC   to the mammal a macromolecular aggregate of the peptide. The
CC   macromolecular aggregate comprises at least 10 peptide subunits, may have
CC   a molecular weight of over 1,000 kDa, and is preferably at least 5 nm in
CC   diameter. It is resistant to digestive degradation, being stabilised in
CC   aggregate form by chemical treatment and/or by recombinant protein
CC   engineering of the peptide. The peptide preferably comprises a HBV
CC   protein selected from HBV surface protein, nucleocapsid protein or
CC   envelope protein. Transmucosal administration to a mammal of a
CC   macromolecular aggregate of a HBV surface protein engenders a systemic
CC   immune response in the mammal. A method of suppressing an immune response
CC   in a mammal involves transmucosally administering a monomolecular peptide
CC   that is resistant to digestive degradation and which may be stabilised by
CC   chemical treatment or protein engineering, and which may be derived from
CC   a HBV protein. A monomolecular peptide is useful for the induction of
CC   oral tolerance when induction of systemic immunity is undesirable, e.g.
CC   in cases of chronic infections.
XX
SQ   Sequence 669 BP; 155 A; 170 C; 148 G; 196 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 30;  DB 12;  Length 669;
  Best Local Similarity   86.7%;  Pred. No. 0.005;
  Matches   26;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||||:|||||||||||:||::||||||||||
Db         73 GACATGAACAAGAGATGATTAGGCAGAGGT 44


Search completed: March 17, 2005, 06:48:41
Job time : 259 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-60
Perfect score: 20
Sequence: 1 gacatgaacaagagatgatt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	30	6	AR027810	AR027810 Sequence
2	20	100.0	30	6	AR027841	AR027841 Sequence
c 3	20	100.0	87	6	AX151115	AX151115 Sequence
c 4	20	100.0	99	14	HPBPRECA	M76687 Hepatitis B
c 5	20	100.0	99	14	HPBPRECB	M76688 Hepatitis B
c 6	20	100.0	99	14	HPBPRECC	M76689 Hepatitis B
c 7	20	100.0	99	14	HPBPRECD	M76690 Hepatitis B
c 8	20	100.0	99	14	HPBPRECE	M76691 Hepatitis B
c 9	20	100.0	99	14	HPBPRECF	M76692 Hepatitis B
c 10	20	100.0	99	14	HPBPRECG	M76693 Hepatitis B
c 11	20	100.0	99	14	HPBPRECH	M76694 Hepatitis B
c 12	20	100.0	99	14	HPBPRECI	M76695 Hepatitis B
c 13	20	100.0	99	14	HPBPRECM	M76699 Hepatitis B
c 14	20	100.0	129	6	AX151114	AX151114 Sequence
c 15	20	100.0	150	14	AF528205	AF528205 Hepatitis
c 16	20	100.0	150	14	AF528206	AF528206 Hepatitis
c 17	20	100.0	150	14	AF528207	AF528207 Hepatitis
c 18	20	100.0	150	14	AF528208	AF528208 Hepatitis
c 19	20	100.0	150	14	AF528209	AF528209 Hepatitis
c 20	20	100.0	150	14	AF528210	AF528210 Hepatitis
c 21	20	100.0	150	14	AF528211	AF528211 Hepatitis
c 22	20	100.0	150	14	AF528212	AF528212 Hepatitis
c 23	20	100.0	150	14	AF528213	AF528213 Hepatitis
c 24	20	100.0	150	14	AF528214	AF528214 Hepatitis
c 25	20	100.0	150	14	AF528215	AF528215 Hepatitis
c 26	20	100.0	150	14	AF528216	AF528216 Hepatitis
c 27	20	100.0	150	14	AF528217	AF528217 Hepatitis
c 28	20	100.0	150	14	AF528218	AF528218 Hepatitis
c 29	20	100.0	150	14	AF528219	AF528219 Hepatitis
c 30	20	100.0	150	14	AF528220	AF528220 Hepatitis
c 31	20	100.0	150	14	AF528221	AF528221 Hepatitis
c 32	20	100.0	150	14	AF528222	AF528222 Hepatitis
c 33	20	100.0	150	14	AF528224	AF528224 Hepatitis
c 34	20	100.0	150	14	AF528225	AF528225 Hepatitis
c 35	20	100.0	150	14	AF528226	AF528226 Hepatitis
c 36	20	100.0	150	14	AF528227	AF528227 Hepatitis
c 37	20	100.0	150	14	AF528228	AF528228 Hepatitis
c 38	20	100.0	150	14	AF528229	AF528229 Hepatitis
c 39	20	100.0	150	14	AF528231	AF528231 Hepatitis
c 40	20	100.0	150	14	AF528232	AF528232 Hepatitis
c 41	20	100.0	150	14	AF528233	AF528233 Hepatitis
c 42	20	100.0	150	14	AF528234	AF528234 Hepatitis
c 43	20	100.0	150	14	AF528235	AF528235 Hepatitis
c 44	20	100.0	150	14	AF528236	AF528236 Hepatitis
c 45	20	100.0	150	14	AF528237	AF528237 Hepatitis

ALIGNMENTS

RESULT 1
AR027810
LOCUS AR027810 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 8 from patent US 5856459.
ACCESSION AR027810
VERSION AR027810.1 GI:5938630
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and Mills,J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 8 05-JAN-1999;
FEATURES Location/Qualifiers
source 1. .30
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATGAACAAGAGATGATT 20
||||||||||||||||||||
Db 1 GACATGAACAAGAGATGATT 20

RESULT 2
AR027841
LOCUS AR027841 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 39 from patent US 5856459.
ACCESSION AR027841
VERSION AR027841.1 GI:5938661
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and

```
            Mills,J.S.
  TITLE     Oligonucleotides specific for hepatitis B virus
  JOURNAL   Patent: US 5856459-A 39 05-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 30;
  Best Local Similarity  100.0%;  Pred. No. 56;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db          1 GACATGAACAAGAGATGATT 20


RESULT 3
AX151115/c
LOCUS       AX151115                 87 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 4 from Patent WO0138498.
ACCESSION   AX151115
VERSION     AX151115.1  GI:14533317
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F.,
            Fried,M. and Rossau,R.
  TITLE     A new genotype of hepatitis b virus
  JOURNAL   Patent: WO 0138498-A 4 31-MAY-2001;
            Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 87;
  Best Local Similarity  100.0%;  Pred. No. 51;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         43 GACATGAACAAGAGATGATT 24


RESULT 4
HPBPRECA/c
LOCUS       HPBPRECA                 99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type1 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76687
VERSION     M76687.1  GI:485341
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45507.1"
                     /db_xref="GI:485342"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 1 (creates internal stop
                     codon)"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity  100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 5
HPBPRECB/c
LOCUS       HPBPRECB                 99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 2precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76688
VERSION     M76688.1  GI:485343
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     variation       2
                     /note="c in wt; t in virus type 2"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45508.1"
                     /db_xref="GI:485344"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 2 (creates internal stop
                     codon)"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity  100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 6
HPBPRECC/c
LOCUS       HPBPRECC                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 3precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76689
VERSION     M76689.1  GI:485345
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     variation       6
                     /note="c in wt; t in virus type 3"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45509.1"
                     /db_xref="GI:485346"
                     /translation="MQLFHLCLIISCSCPTFQASKLCLGWL"
     variation       58
                     /gene="C"
                     /note="g in wt; t in virus type 3 (val to phe)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 3 (creates internal stop
                     codon)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 7
HPBPRECD/c
LOCUS       HPBPRECD                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 4 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76690
VERSION     M76690.1  GI:485347
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45510.1"
                     /db_xref="GI:485348"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 4 (creates internal stop
                     codon)"
     variation       95
                     /note="g in wt; a in virus type 4 (gly to asp)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 8
HPBPRECE/c
LOCUS       HPBPRECE                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 5 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76691
VERSION     M76691.1  GI:485349
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45511.1"
```

```
                     /db_xref="GI:485350"
                     /translation="MQLFHLCLIISCSCPTVQPSKLCLGWL"
     variation       64. .67
                     /gene="C"
                     /note="gcc in wt; ccg in virus type 5 (ala to pro)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 5 (creates internal stop
                     codon)"
     variation       95
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 9
HPBPRECF/c
LOCUS       HPBPRECF                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 6 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76692
VERSION     M76692.1  GI:485351
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .99
                     /gene="C"
     misc_feature    10. .99
                     /gene="C"
                     /product="precore protein"
                     /note="putative cds"
     variation       11
                     /gene="C"
                     /note="t in wt; c in virus type 6 (loss of start codon)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 10
HPBPRECG/c
LOCUS       HPBPRECG                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 7 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76693
VERSION     M76693.1  GI:485352
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     misc_feature    10. .93
                     /gene="C"
                     /product="precore protein"
                     /standard_name="pre-C region note: putative CDS"
     variation       10
                     /gene="C"
                     /note="a in wt; t in virus type 7 (loss of start codon)"
     variation       14
                     /gene="C"
                     /note="a in wt; g in virus type 7 (gln to arg)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 7 ( creates internal stop
                     codon)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 11
HPBPRECH/c
LOCUS       HPBPRECH                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 8 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76694
VERSION     M76694.1  GI:485353
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     misc_feature    10. .93
```

```
                     /gene="C"
                     /product="precore protein"
                     /standard_name="pre-C region note: putative CDS"
     variation       12
                     /gene="C"
                     /note="g in wt; t in virus type 8 (loss of start codon)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 8 (creates internal stop
                     codon)"
     variation       95
                     /note="g in wt; a in virus type 8 (gly to asp)"
ORIGIN

  Query Match            100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity  100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 12
HPBPRECI/c
LOCUS       HPBPRECI                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 9 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76695
VERSION     M76695.1  GI:485354
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     misc_feature    10. .93
                     /gene="C"
                     /product="precore protein"
                     /standard_name="pre-C region note: putative CDS"
     variation       13
                     /gene="C"
                     /note="c in wt; t in virus type 9 (creates internal stop
                     codon)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 9 (creates internal stop
                     codon)"
     variation       95
                     /note="g in wt; a in virus type 9 (gly to asp)"
ORIGIN

  Query Match            100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity  100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 13
HPBPRECM/c
LOCUS       HPBPRECM                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 13 precore protein (pre-C region, C) gene,
            5' end.
ACCESSION   M76699
VERSION     M76699.1  GI:485361
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .99
                     /gene="C"
     CDS             10. .>99
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45515.1"
                     /db_xref="GI:485362"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWLWDM"
     variation       95
                     /gene="C"
                     /note="g in wt; a in virus type 13 (gly to asp)"
ORIGIN

  Query Match            100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity  100.0%;  Pred. No. 50;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         52 GACATGAACAAGAGATGATT 33


RESULT 14
AX151114/c
LOCUS       AX151114                 129 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3 from Patent WO0138498.
ACCESSION   AX151114
VERSION     AX151114.1  GI:14533316
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F.,
            Fried,M. and Rossau,R.
  TITLE     A new genotype of hepatitis b virus
  JOURNAL   Patent: WO 0138498-A 3 31-MAY-2001;
            Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .129
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN
```

```
  Query Match             100.0%;  Score 20;  DB 6;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 49;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         43 GACATGAACAAGAGATGATT 24


RESULT 15
AF528205/c
LOCUS       AF528205                150 bp    DNA     linear   VRL 31-JUL-2003
DEFINITION  Hepatitis B virus ASC1123 core antigen precursor, gene, partial
            cds.
ACCESSION   AF528205
VERSION     AF528205.1  GI:32810971
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 150)
  AUTHORS   Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
  TITLE     Comparative evaluation of HBV precore and basal core promoter
            mutants in Indian patients with diverse clinical manifestations
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150)
  AUTHORS   Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUL-2002) Hepatitis Division, National Institute of
            Virology, 20-A, Dr Ambedkar Road, Pune, Maharashtra 411001, India
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="Hepatitis B virus"
                     /proviral
                     /mol_type="genomic DNA"
                     /isolate="ASC1123"
                     /isolation_source="asymptomatic HBsAg carrier"
                     /specific_host="Homo sapiens"
                     /db_xref="taxon:10407"
                     /country="India"
     misc_feature    <1. .>150
                     /note="contains partial basal core promoter"
     CDS             64. .>150
                     /note="contains complete precore region"
                     /codon_start=1
                     /product="core antigen precursor"
                     /protein_id="AAP87556.1"
                     /db_xref="GI:32810972"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWLXG"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 150;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db        106 GACATGAACAAGAGATGATT 87


Search completed: March 17, 2005, 08:14:16
Job time : 684.733 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 2079.4 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-59
Perfect score: 30
Sequence: 1 gacaugaacaagagaugauuaggcagaggt 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	22	73.3	661	9	CG452180	CG452180 OG5DV84TC
	2	22	73.3	663	6	CD304890	CD304890 StrPu691.
c	3	22	73.3	701	8	AQ993161	AQ993161 RPCI-23-3
	4	22	73.3	888	9	CG359992	CG359992 OG1CX42TV
c	5	21.2	70.7	577	4	BI378081	BI378081 BFLG3_001
	6	21.2	70.7	843	9	CW009369	CW009369 ZMMBLa001
c	7	21	70.0	214	2	AW342249	AW342249 GthEST1 G
	8	21	70.0	345	5	BY143963	BY143963 BY143963
	9	21	70.0	399	6	CD296187	CD296187 StrPu691.
	10	21	70.0	519	2	BF103720	BF103720 601647304
c	11	21	70.0	693	9	CR048521	CR048521 Forward s
c	12	21	70.0	880	4	BG104640	BG104640 602311331
	13	21	70.0	1110	8	CC258547	CC258547 CH261-63K
	14	21	70.0	1154	2	BF342749	BF342749 602015013
	15	20.8	69.3	73	9	CG623002	CG623002 OST323776
c	16	20.8	69.3	715	8	AQ939213	AQ939213 NL1-I20R
	17	20.6	68.7	223	9	CE129305	CE129305 tigr-gss-
c	18	20.6	68.7	370	9	CL423226	CL423226 RP11-413M
c	19	20.6	68.7	558	8	AQ272786	AQ272786 nbxb0028N
	20	20.6	68.7	577	6	CA044022	CA044022 ssalpha50
c	21	20.6	68.7	595	8	AQ509958	AQ509958 nbxb0094H
c	22	20.6	68.7	600	8	AQ257439	AQ257439 nbxb0018A
	23	20.6	68.7	622	9	CL718440	CL718440 OR_BBa004
c	24	20.6	68.7	655	6	CA061518	CA061518 ssalrgb52
c	25	20.6	68.7	663	6	CA061593	CA061593 ssalrga50
	26	20.6	68.7	680	9	CE792348	CE792348 tigr-gss-
c	27	20.6	68.7	699	9	CL779656	CL779656 OR_BBa009
c	28	20.6	68.7	705	9	CL612489	CL612489 OR_BBa000
c	29	20.6	68.7	715	1	AV707321	AV707321 AV707321
	30	20.6	68.7	733	6	CA348548	CA348548 679903 NC
c	31	20.6	68.7	757	6	CB512505	CB512505 ssalrgb55
c	32	20.6	68.7	767	9	CL787302	CL787302 OR_BBa010
	33	20.6	68.7	776	9	CL793261	CL793261 OR_CBa000
	34	20.6	68.7	780	8	AQ687676	AQ687676 nbxb0075K
c	35	20.6	68.7	854	9	CL795294	CL795294 OR_CBa000
	36	20.6	68.7	1012	7	CN247122	CN247122 EST013009
	37	20.4	68.0	344	8	BZ798022	BZ798022 PUGDA69TD
	38	20.4	68.0	447	8	BZ642588	BZ642588 OGAMX61TC
	39	20.4	68.0	490	6	CD339264	CD339264 StrPu536.
	40	20.4	68.0	501	6	CD330453	CD330453 StrPu537.
	41	20.4	68.0	510	1	AA003912	AA003912 mg79e03.r
	42	20.4	68.0	572	6	CD309830	CD309830 StrPu691.
c	43	20.4	68.0	586	9	FR0003872	Z87655 F.rubripes
	44	20.4	68.0	678	6	CD310469	CD310469 StrPu691.
	45	20.4	68.0	727	9	CL557140	CL557140 OB__Ba000

ALIGNMENTS

RESULT 1
CG452180/c
LOCUS CG452180 661 bp DNA linear GSS 17-SEP-2003
DEFINITION OG5DV84TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0841N24, genomic survey sequence.
ACCESSION CG452180
VERSION CG452180.1 GI:34837180
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 661)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .661
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0841N24"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 73.3%; Score 22; DB 9; Length 661;
Best Local Similarity 76.7%; Pred. No. 1.6e+02;
Matches 23; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              | ||:||| ||||||:|| | ||||||||
Db        481 GTCATGAAGAAGAGATGAAGAAGCAGAGGT 452
```

```
RESULT 2
CD304890
LOCUS       CD304890                663 bp    mRNA    linear   EST 16-SEP-2003
DEFINITION  StrPu691.001255 Sea urchin larva cDNA library MPMGp691
            Strongylocentrotus purpuratus cDNA clone
            MPMGp691C0520;MPI_SURUDI_20C5 5', mRNA sequence.
ACCESSION   CD304890
VERSION     CD304890.1  GI:34749939
KEYWORDS    EST.
SOURCE      Strongylocentrotus purpuratus
  ORGANISM  Strongylocentrotus purpuratus
            Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
            Echinoidea; Euechinoidea; Echinacea; Echinoida;
            Strongylocentrotidae; Strongylocentrotus.
REFERENCE   1  (bases 1 to 663)
  AUTHORS   Poustka,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,
            Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.
  TITLE     Generation, annotation, evolutionary analysis, and database
            integration of 20,000 unique sea urchin EST clusters
  JOURNAL   Genome Res. 13 (12), 2736-2746 (2003)
COMMENT     Contact: Poustka AJ
            laboraty 145, dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1235
            Fax: +49 30 8413 1128
            Email: poustka@molgen.mpg.de
            The library was characterised by oligonucleotide fingerprinting
            (ONF) to reduce sequencing redundancy. According to the ONF
            procedure, clones that display the same hybridisation matrix with a
            battery of 200 8mer oligonucleotides are grouped into clusters. One
            clone per ONF cluster is selected for sequencing. The size of each
            cluster is an indicator of the frequency of a transcript in the
            analysed library. The cluster size as well as the coordinates of
            the other clones assigned to the same ONF cluster as the clone from
            which the above EST is generated is available at the sea urchin
            project web site at: http://www.molgen.mpg.de/ag_seaurchin/ . cDNA
            clones and filters are distributed via the Resource Center/Primary
            Database of the German Human Genome Project (http://www.rzpd.de)
            PCR PRimers
            FORWARD: 5' CCCCAGGCTTTACACTTTATGCTTCCGGCTCG 3' (M13RSP) 5'-seq
            BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGGATGTG 3' (M13FSP) 3'-seq
            Seq primer: 5'-CCGGTCCGGAATTCCCGGGT-3' pSport3/86
            High quality sequence stop: 663.
FEATURES             Location/Qualifiers
     source          1. .663
                     /organism="Strongylocentrotus purpuratus"
                     /mol_type="mRNA"
                     /db_xref="taxon:7668"
                     /clone="MPMGp691C0520;MPI_SURUDI_20C5"
                     /tissue_type="whole larva"
                     /dev_stage="larva 2-3 weeks"
                     /lab_host="E.coli, XL1 blue"
                     /clone_lib="Sea urchin larva cDNA library MPMGp691"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI; Random
                     primed and directionally cloned in pSport1 vector using a
                     NotI (5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC (T)15-3' and a
                     SalI 5'- TCGACCCACGCGTCCG-3'adapters (Gibco BRL)"
ORIGIN

  Query Match             73.3%;  Score 22;  DB 6;  Length 663;
  Best Local Similarity   76.7%;  Pred. No. 1.6e+02;
  Matches   23;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              |||| |||  ||||| ||::||| ||||||
Db        123 GACAAGAAGGAGAGAAGATTAGGAAGAGGT 152


RESULT 3
AQ993161/c
LOCUS       AQ993161                701 bp    DNA     linear   GSS 24-FEB-2000
DEFINITION  RPCI-23-370H16.TV RPCI-23 Mus musculus genomic clone
            RPCI-23-370H16, genomic survey sequence.
ACCESSION   AQ993161
VERSION     AQ993161.1  GI:7068258
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 701)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-370H16.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 370  row: H  column: 16
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .701
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-370H16"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

  Query Match             73.3%;  Score 22;  DB 8;  Length 701;
  Best Local Similarity   73.3%;  Pred. No. 1.7e+02;
  Matches   22;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              ||  :|||||||||| ||::|||  |||||
Db        519 GAAGTGAACAAGAGAAGATTAGGGGGAGGT 490


RESULT 4
CG359992
LOCUS       CG359992                888 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OG1CX42TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0733H12,
            genomic survey sequence.
ACCESSION   CG359992
VERSION     CG359992.1  GI:34277259
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
```

```
REFERENCE   1  (bases 1 to 888)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1CX42TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .888
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0733H12"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             73.3%;  Score 22;  DB 9;  Length 888;
  Best Local Similarity   76.7%;  Pred. No. 1.7e+02;
  Matches   23;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              | ||:||| |||||||:||  | ||||||||
Db        245 GTCATGAAGAAGAGATGAAGAAGCAGAGGT 274


RESULT 5
BI378081/c
LOCUS       BI378081                577 bp    mRNA    linear   EST 26-AUG-2003
DEFINITION  BFLG3_001884 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG
            or MPMGp498) Branchiostoma floridae cDNA clone MPMGp498N0628 5',
            mRNA sequence.
ACCESSION   BI378081
VERSION     BI378081.1  GI:30913195
KEYWORDS    EST.
SOURCE      Branchiostoma floridae (Florida lancelet)
  ORGANISM  Branchiostoma floridae
            Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
            Branchiostoma.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J.,
            Herwig,R., Vingron,M. and Lehrach,H.
  TITLE     New evidence for genome-wide duplications at the origin of
            vertebrates using an amphioxus gene set and completed animal
            genomes
  JOURNAL   Genome Res. 13 (6A), 1056-1066 (2003)
  MEDLINE   22683279
   PUBMED   12799346
COMMENT     Contact: Panopoulou G
            laboratory 145, dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1235
            Fax: +49 30 8413 1128
            Email: panopoul@molgen.mpg.de
            The library was characterised by oligonucleotide fingerprinting
            (ONFP) to reduce sequencing redundancy. According to the ONFP
            procedure, clones giving the same hybridisation pattern with a
            battery of 200 8mer oligonucleotides are grouped into clusters. One
            clone per cluster is selected for sequencing. The size of each
            cluster is an indicator of the frequency of a transcript in the
            analysed library. The cluster size as well the coordinates of the
            rest of the clones assigned to the same fingerprint cluster as the
            clone from which the above EST is generated is available at the
            amphioxus project site at: http://www.molgen.mpg.de/amphioxus/
            Clones and filters are distributed via the Resource Center/Primary
            Database of the German Genome Project (http://www.rzpd.de)
            PCR PRimers
            FORWARD: 5' CCCCAGGCTTTACACTTTATGCTTCCGGCTCG 3' (M13RSP)
            BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGGATGTG 3' (M13FSP)
            Insert Length: 1200    Std Error: 0.00
            Seq primer: 5'-CCGGTCCGGAATTCCCGGGT-3' pSport3/86
            High quality sequence stop: 577.
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Branchiostoma floridae"
                     /mol_type="mRNA"
                     /db_xref="taxon:7739"
                     /clone="MPMGp498N0628"
                     /tissue_type="whole embryo"
                     /dev_stage="5-6 hrs (gastrula stage)"
                     /lab_host="E.coli, XL1 blue"
                     /clone_lib="Amphioxus 5-6 hrs cDNA library (Name
                     convention: BFLG or MPMGp498)"
                     /note="Vector: pSport1; Site_1: SalI, KpnI, EcoRI (5');
                     Site_2: NotI, BamHI, HindIII (3'); OligodT primed and
                     directionally cloned in pSport1 vector using a  NotI
                     (5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC (T)15-3' and a SalI 5'-
                     TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."
ORIGIN

  Query Match             70.7%;  Score 21.2;  DB 4;  Length 577;
  Best Local Similarity   73.1%;  Pred. No. 3.5e+02;
  Matches   19;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 AUGAACAAGAGAUGAUUAGGCAGAGG 29
              |:||||||||||:||:: | |||| |
Db        507 ATGAACAAGAGATGATTGGTCAGAAG 482


RESULT 6
CW009369
LOCUS       CW009369                843 bp    DNA     linear   GSS 23-SEP-2004
DEFINITION  ZMMBLa0012A22.f ZMMBLa Zea mays genomic clone ZMMBLa0012A22 5',
            genomic survey sequence.
ACCESSION   CW009369
VERSION     CW009369.1  GI:52591950
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 843)
  AUTHORS   Wing,R., Luo,M., Soderlund,C. and Haller,K.
  TITLE     ZMMBL sequences
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: http://genome.arizona.edu
            Plate: 0012  row: A  column: 22
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .843
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBLa0012A22"
                     /tissue_type="immature ears"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="ZMMBLa"
```

/note="Vector: pAGIBAC1; Site_1: SalI; Site_2: SalI"
ORIGIN

Query Match 70.7%; Score 21.2; DB 9; Length 843;
Best Local Similarity 73.1%; Pred. No. 3.7e+02;
Matches 19; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAGAGAUGAUUAGGCAG 26
|| |:|||||||||:| ::|||| |
Db 817 GAGATGAACAAGAGATGTTTAGGCTG 842

RESULT 7
AW342249/c
LOCUS AW342249 214 bp mRNA linear EST 01-JAN-2001
DEFINITION GthEST1 Guillardia theta Lambda Zap II cDNA Library Guillardia theta cDNA clone 38r, mRNA sequence.
ACCESSION AW342249
VERSION AW342249.1 GI:12000490
KEYWORDS EST.
SOURCE Guillardia theta
ORGANISM Guillardia theta
Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
REFERENCE 1 (bases 1 to 214)
AUTHORS Fraunholz,M., Duebel,J., Wastl,J., Zauner,S. and Maier,U.-G.
TITLE EST Database of the cryptomonad alga: Guillardia theta
JOURNAL Unpublished (2000)
COMMENT Contact: Maier, U.-G.
Department of Cell Biology and Applied Botany
Philipps-University Marburg
Karl-von-Frisch-Strasse, D-35043 Marburg, Germany
Tel: ++49 6421 282 2057
Fax: ++49 6421 282 1543
Email: maier@mailer.uni-marburg.de.
FEATURES Location/Qualifiers
source 1. .214
/organism="Guillardia theta"
/mol_type="mRNA"
/strain="CCMP327"
/db_xref="taxon:55529"
/clone="38r"
/lab_host="SOLR"
/clone_lib="Guillardia theta Lambda Zap II cDNA Library"
/note="Vector: Lambda ZAP II; Site_1: EcoRI"
ORIGIN

Query Match 70.0%; Score 21; DB 2; Length 214;
Best Local Similarity 72.4%; Pred. No. 3.6e+02;
Matches 21; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAGAGAUGAUUAGGCAGAGG 29
|| | | || ||||:||::|||||||||
Db 89 GAAAAAAGCAGGAGATGATTAGGCAGAGG 61

RESULT 8
BY143963
LOCUS BY143963 345 bp mRNA linear EST 09-DEC-2002
DEFINITION BY143963 RIKEN full-length enriched, 17.5 days embryo whole body Mus musculus cDNA clone L930183C02 5', mRNA sequence.
ACCESSION BY143963
VERSION BY143963.1 GI:26280016
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 345)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 22354683
PUBMED 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .345
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L930183C02"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"
ORIGIN

Query Match 70.0%; Score 21; DB 5; Length 345;

```
  Best Local Similarity   75.9%;  Pred. No. 3.9e+02;
  Matches   22;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGG 29
              ||||:| ||||||||  |: || ||||||
Db         72 GACATGCACAAGAGAGAATGAGCCAGAGG 100


RESULT 9
CD296187
LOCUS       CD296187                399 bp    mRNA    linear   EST 16-SEP-2003
DEFINITION  StrPu691.007582 Sea urchin larva cDNA library MPMGp691
            Strongylocentrotus purpuratus cDNA clone
            MPMGp691G19122;MPI_SURUDI_122G19 5', mRNA sequence.
ACCESSION   CD296187
VERSION     CD296187.1  GI:34747264
KEYWORDS    EST.
SOURCE      Strongylocentrotus purpuratus
  ORGANISM  Strongylocentrotus purpuratus
            Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
            Echinoidea; Euechinoidea; Echinacea; Echinoida;
            Strongylocentrotidae; Strongylocentrotus.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Poustka,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,
            Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.
  TITLE     Generation, annotation, evolutionary analysis, and database
            integration of 20,000 unique sea urchin EST clusters
  JOURNAL   Genome Res. 13 (12), 2736-2746 (2003)
COMMENT     Contact: Poustka AJ
            laboraty 145, dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1235
            Fax: +49 30 8413 1128
            Email: poustka@molgen.mpg.de
            The library was characterised by oligonucleotide fingerprinting
            (ONF) to reduce sequencing redundancy. According to the ONF
            procedure, clones that display the same hybridisation matrix with a
            battery of 200 8mer oligonucleotides are grouped into clusters. One
            clone per ONF cluster is selected for sequencing. The size of each
            cluster is an indicator of the frequency of a transcript in the
            analysed library. The cluster size as well as the coordinates of
            the other clones assigned to the same ONF cluster as the clone from
            which the above EST is generated is available at the sea urchin
            project web site at: http://www.molgen.mpg.de/ag_seaurchin/ . cDNA
            clones and filters are distributed via the Resource Center/Primary
            Database of the German Human Genome Project (http://www.rzpd.de)
            PCR PRimers
            FORWARD: 5' CCCCAGGCTTTACACTTTATGCTTCCGGCTCG 3' (M13RSP) 5'-seq
            BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGGATGTG 3' (M13FSP) 3'-seq
            Seq primer: 5'-CCGGTCCGGAATTCCCGGGT-3' pSport3/86
            High quality sequence stop: 399.
FEATURES             Location/Qualifiers
     source          1. .399
                     /organism="Strongylocentrotus purpuratus"
                     /mol_type="mRNA"
                     /db_xref="taxon:7668"
                     /clone="MPMGp691G19122;MPI_SURUDI_122G19"
                     /tissue_type="whole larva"
                     /dev_stage="larva 2-3 weeks"
                     /lab_host="E.coli, XL1 blue"
                     /clone_lib="Sea urchin larva cDNA library MPMGp691"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI; Random
                     primed and directionally cloned in pSport1 vector using a
                     NotI (5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC (T)15-3' and a
                     SalI 5'- TCGACCCACGCGTCCG-3'adapters (Gibco BRL)"
ORIGIN

  Query Match             70.0%;  Score 21;  DB 6;  Length 399;
  Best Local Similarity   73.3%;  Pred. No. 4e+02;
  Matches   22;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
                ||||  |||  ||||| ||::||| ||| ||
Db         99 GACANGAAGGAGAGAAGATTAGGAAGAAGT 128


RESULT 10
BF103720
LOCUS       BF103720                519 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  601647304F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:3931440 5',
            mRNA sequence.
ACCESSION   BF103720
VERSION     BF103720.1  GI:10886246
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM767  row: h  column: 01
            High quality sequence stop: 518.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3931440"
                     /tissue_type="embryonal carcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_61"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             70.0%;  Score 21;  DB 2;  Length 519;
  Best Local Similarity   75.9%;  Pred. No. 4.2e+02;
  Matches   22;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGG 29
              || | |||||||||  ||::| |||||||
Db        488 GAAAGGAACAAGAGCAGATTAAGCAGAGG 516


RESULT 11
CR048521/c
LOCUS       CR048521                693 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHPP323e21, genomic survey sequence.
ACCESSION   CR048521
VERSION     CR048521.1  GI:49781660
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .693
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPP323e21"
                     /clone_lib="MHPP"
ORIGIN

  Query Match             70.0%;  Score 21;  DB 9;  Length 693;
  Best Local Similarity   72.4%;  Pred. No. 4.4e+02;
  Matches   21;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ACAUGAACAAGAGAUGAUUAGGCAGAGGT 30
              |||:||| ||||||:|  :||| | ||||
Db        454 ACATGAAAAAGAGATGTCTAGGAAAAGGT 426


RESULT 12
BG104640/c
LOCUS       BG104640                880 bp    mRNA    linear   EST 30-JAN-2001
DEFINITION  602311331F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4421191 5',
            mRNA sequence.
ACCESSION   BG104640
VERSION     BG104640.1  GI:12598486
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 880)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10159  row: n  column: 08
            High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1. .880
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4421191"
                     /tissue_type="adrenal cortex carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_84"
                     /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
                     NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
                     primed. Average insert size 1.229 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             70.0%;  Score 21;  DB 4;  Length 880;
  Best Local Similarity   79.3%;  Pred. No. 4.5e+02;
  Matches   23;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGG 29
              || | |||||||||| :||  |||||||||
Db         98 GAAAGGAACAAGAGGTGAGAAGGCAGAGG 70


RESULT 13
CC258547
LOCUS       CC258547                1110 bp    DNA     linear   GSS 13-MAY-2003
DEFINITION  CH261-63K24_RM1.1 CH261 Gallus gallus genomic clone CH261-63K24,
            genomic survey sequence.
ACCESSION   CC258547
VERSION     CC258547.1  GI:30599491
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1110)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 182000  Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 43
            High quality sequence stop: 696.
FEATURES             Location/Qualifiers
     source          1. .1110
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-63K24"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
ORIGIN

  Query Match             70.0%;  Score 21;  DB 8;  Length 1110;
  Best Local Similarity   72.4%;  Pred. No. 4.7e+02;
  Matches   21;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGG 29
              || |:||| ||||||: | :||||| |||
Db        407 GAAATGAAGAAGAGATAAATAGGCACAGG 435


RESULT 14
BF342749
LOCUS       BF342749                1154 bp    mRNA    linear   EST 22-NOV-2000
DEFINITION  602015013F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4150755
            5', mRNA sequence.
ACCESSION   BF342749
VERSION     BF342749.1  GI:11289773
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1154)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9414  row: j  column: 04
            High quality sequence stop: 657.
FEATURES              Location/Qualifiers
     source           1. .1154
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4150755"
                      /tissue_type="glioblastoma with EGFR amplification"
                      /lab_host="DH10B (T1 phage-resistant)"
                      /clone_lib="NCI_CGAP_Brn64"
                      /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                      Average insert size 1.57 kb. Constructed by Life
                      Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match           70.0%;  Score 21;  DB 2;  Length 1154;
  Best Local Similarity   69.0%;  Pred. No. 4.7e+02;
  Matches   20;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUUAGGCAGAGG 29
              ||| :| || |||||: |:: ||||||||
Db        328 GACCTGGACGAGAGATTATTTGGCAGAGG 356


RESULT 15
CG623002
LOCUS       CG623002                  73 bp    mRNA    linear   GSS 02-OCT-2003
DEFINITION  OST323776 Mus musculus 129Sv/Ev Mus musculus cDNA clone OST323776,
            mRNA sequence.
ACCESSION   CG623002
VERSION     CG623002.1  GI:37446851
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES              Location/Qualifiers
     source           1. .73
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="129Sv/Ev"
                      /db_xref="taxon:10090"
                      /clone="OST323776"
                      /cell_type="embryonic stem cell"
                      /clone_lib="Mus musculus 129Sv/Ev"
ORIGIN

  Query Match           69.3%;  Score 20.8;  DB 9;  Length 73;
  Best Local Similarity   66.7%;  Pred. No. 3.7e+02;
  Matches   18;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ACAUGAACAAGAGAUGAUUAGGCAGAG 28
              |||:| || |||||: |:: ||||| |
Db         19 ACATGNACCAGAGATNATTNGGCAGGG 45


Search completed: March 17, 2005, 11:07:42
Job time : 2082.4 secs
```

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-60
Perfect score: 20
Sequence: 1 gacatgaacaagagatgatt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	2	AAT72566	Aat72566 Hepatitis
	2	20	100.0	20	2	AAT72565	Aat72565 Hepatitis
	3	20	100.0	30	2	AAT72562	Aat72562 Hepatitis
	4	20	100.0	30	2	AAT72563	Aat72563 Hepatitis
	5	20	100.0	30	2	AAT72616	Aat72616 Hepatitis
	6	20	100.0	30	2	AAT72617	Aat72617 Hepatitis
c	7	20	100.0	39	10	ADC64742	Adc64742 Hepatitis
c	8	20	100.0	87	4	AAD09094	Aad09094 Hepatitis
c	9	20	100.0	129	4	AAD09093	Aad09093 Hepatitis
c	10	20	100.0	639	6	AAD27422	Aad27422 Hepatitis
c	11	20	100.0	639	6	AAD31509	Aad31509 Hepatitis
c	12	20	100.0	655	4	AAH77569	Aah77569 HBV genot
c	13	20	100.0	655	4	AAH77568	Aah77568 HBV genot
c	14	20	100.0	655	4	AAH77574	Aah77574 HBV genot
c	15	20	100.0	655	4	AAH77573	Aah77573 HBV genot
c	16	20	100.0	655	4	AAH77570	Aah77570 HBV genot
c	17	20	100.0	655	4	AAH77571	Aah77571 HBV genot
c	18	20	100.0	664	4	AAH77572	Aah77572 HBV genot
c	19	20	100.0	669	12	ADO07220	Ado07220 Hepatitis
c	20	20	100.0	673	4	AAD09092	Aad09092 Hepatitis
c	21	20	100.0	675	4	AAH77563	Aah77563 HBV preCo
c	22	20	100.0	681	4	AAH77567	Aah77567 HBV genot
c	23	20	100.0	1395	2	AAV82688	Aav82688 Fulminant
c	24	20	100.0	1400	2	AAV82687	Aav82687 Fulminant
c	25	20	100.0	1445	2	AAV82692	Aav82692 Fulminant
c	26	20	100.0	1445	2	AAV82685	Aav82685 Fulminant
c	27	20	100.0	1445	2	AAV82690	Aav82690 Fulminant
c	28	20	100.0	1445	2	AAV82684	Aav82684 Fulminant
c	29	20	100.0	1500	2	AAV82695	Aav82695 Fulminant
c	30	20	100.0	1500	2	AAV82683	Aav82683 Fulminant
c	31	20	100.0	1500	2	AAV82694	Aav82694 Fulminant
c	32	20	100.0	1500	2	AAV82686	Aav82686 Fulminant
c	33	20	100.0	1500	2	AAV82706	Aav82706 Wild type
c	34	20	100.0	1500	2	AAV82689	Aav82689 Fulminant
c	35	20	100.0	1500	2	AAV82693	Aav82693 Fulminant
c	36	20	100.0	2342	1	AAN93072	Aan93072 Sequence
c	37	20	100.0	2743	1	AAN00003	Aan00003 Sequence
c	38	20	100.0	2743	2	AAQ04799	Aaq04799 Recombina
c	39	20	100.0	3180	4	AAH42375	Aah42375 Nucleotid
c	40	20	100.0	3182	6	AAD31765	Aad31765 Hepatitis
c	41	20	100.0	3182	9	ACA62422	Aca62422 Hepatitis
c	42	20	100.0	3182	10	AAD60866	Aad60866 Hepatitis
c	43	20	100.0	3220	3	AAZ88924	Aaz88924 Hepatitis
c	44	20	100.0	3248	4	AAD09091	Aad09091 Hepatitis
c	45	20	100.0	3248	4	AAH77562	Aah77562 HBV genot

ALIGNMENTS

RESULT 1
AAT72566
ID AAT72566 standard; RNA; 20 BP.
XX
AC AAT72566;
XX
DT 03-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV46MYb.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1. .20
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
FT modified_base 1
FT /*tag= b
FT /mod_base= gm
FT modified_base 2
FT /*tag= c
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 3
FT /*tag= d
FT /mod_base= cm
FT modified_base 4
FT /*tag= e
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 5
FT /*tag= f
FT /mod_base= um
FT modified_base 6
FT /*tag= g
FT /mod_base= gm
FT modified_base 7
FT /*tag= h
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 8

```
FT                   /*tag=  i
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   9
FT                   /*tag=  j
FT                   /mod_base= cm
FT   modified_base   10
FT                   /*tag=  k
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   11
FT                   /*tag=  l
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   12
FT                   /*tag=  m
FT                   /mod_base= gm
FT   modified_base   13
FT                   /*tag=  n
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   14
FT                   /*tag=  o
FT                   /mod_base= gm
FT   modified_base   15
FT                   /*tag=  p
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   16
FT                   /*tag=  q
FT                   /mod_base= um
FT   modified_base   17
FT                   /*tag=  r
FT                   /mod_base= gm
FT   modified_base   18
FT                   /*tag=  s
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   19
FT                   /*tag=  t
FT                   /mod_base= um
FT   modified_base   20
FT                   /*tag=  u
FT                   /mod_base= um
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV46MYb
CC   which is complementary to a portion of the hepatitis B virus (HBV) RNA.
CC   The antisense oligonucleotide may be used to detect the presence of HBV
CC   in a sample. The antisense oligonucleotide, and oligonucleotides
CC   containing a sequence which is complementary to at least two non-
CC   contiguous regions of an HBV nucleic acid, may be used for inhibiting HBV
CC   replication in a cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 9 A; 2 C; 5 G; 0 T; 4 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 9.6;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||:||||||||||:||::
Db          1 GACAUGAACAAGAGAUGAUU 20


RESULT 2
AAT72565
ID   AAT72565 standard; DNA; 20 BP.
XX
AC   AAT72565;
XX
DT   03-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV46Yb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .20
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV46Yb which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 9 A; 2 C; 5 G; 4 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 9.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db          1 GACATGAACAAGAGATGATT 20


RESULT 3
AAT72562
```

```
ID   AAT72562 standard; DNA; 30 BP.
XX
AC   AAT72562;
XX
DT   03-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV88b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag= a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV88b which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 10 G; 5 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db          1 GACATGAACAAGAGATGATT 20


RESULT 4
AAT72563
ID   AAT72563 standard; DNA; 30 BP.
XX
AC   AAT72563;
XX
DT   03-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV88Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag= a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        1. .20
FT                   /*tag= b
FT                   /note= "2'-OMe RNA"
FT   modified_base   1
FT                   /*tag= c
FT                   /mod_base= gm
FT   modified_base   2
FT                   /*tag= d
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   3
FT                   /*tag= e
FT                   /mod_base= cm
FT   modified_base   4
FT                   /*tag= f
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   5
FT                   /*tag= g
FT                   /mod_base= um
FT   modified_base   6
FT                   /*tag= h
FT                   /mod_base= gm
FT   modified_base   7
FT                   /*tag= i
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   8
FT                   /*tag= j
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   9
FT                   /*tag= k
FT                   /mod_base= cm
FT   modified_base   10
FT                   /*tag= l
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   11
FT                   /*tag= m
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   12
FT                   /*tag= n
FT                   /mod_base= gm
FT   modified_base   13
FT                   /*tag= o
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   14
FT                   /*tag= p
FT                   /mod_base= gm
FT   modified_base   15
FT                   /*tag= q
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   16
FT                   /*tag= r
FT                   /mod_base= um
FT   modified_base   17
FT                   /*tag= s
FT                   /mod_base= gm
FT   modified_base   18
FT                   /*tag= s
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   19
FT                   /*tag= u
FT                   /mod_base= um
FT   modified_base   20
FT                   /*tag= v
```

```
FT                   /mod_base= um
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV88Mb which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 10 G; 1 T; 4 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 9.9;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||:|||||||||:||::
Db          1 GACAUGAACAAGAGAUGAUU 20


RESULT 5
AAT72616
ID   AAT72616 standard; DNA; 30 BP.
XX
AC   AAT72616;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-89b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-89b which
CC   contains a sequence which is complementary to at least two non-contiguous
CC   regions of a hepatitis B virus (HBV) nucleic acid. The antisense
CC   oligonucleotide may be used to detect the presence of HBV in a sample.
CC   The antisense oligonucleotide, and oligonucleotides complementary to a
CC   portion of the HBV RNA, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 9 G; 6 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db          1 GACATGAACAAGAGATGATT 20


RESULT 6
AAT72617
ID   AAT72617 standard; DNA; 30 BP.
XX
AC   AAT72617;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-89Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        1. .20
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   1
FT                   /*tag=  c
FT                   /mod_base= gm
FT   modified_base   2
FT                   /*tag=  d
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   3
FT                   /*tag=  e
FT                   /mod_base= cm
FT   modified_base   4
FT                   /*tag=  f
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   5
FT                   /*tag=  g
FT                   /mod_base= um
FT   modified_base   6
FT                   /*tag=  h
FT                   /mod_base= gm
FT   modified_base   7
FT                   /*tag=  i
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
```

```
FT   modified_base   8
FT                   /*tag= j
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   9
FT                   /*tag= k
FT                   /mod_base= cm
FT   modified_base   10
FT                   /*tag= l
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   11
FT                   /*tag= m
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   12
FT                   /*tag= n
FT                   /mod_base= gm
FT   modified_base   13
FT                   /*tag= o
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   14
FT                   /*tag= p
FT                   /mod_base= gm
FT   modified_base   15
FT                   /*tag= q
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   16
FT                   /*tag= r
FT                   /mod_base= um
FT   modified_base   17
FT                   /*tag= s
FT                   /mod_base= gm
FT   modified_base   18
FT                   /*tag= t
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   19
FT                   /*tag= u
FT                   /mod_base= um
FT   modified_base   20
FT                   /*tag= v
FT                   /mod_base= um
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-89Mb
CC   which contains a sequence which is complementary to at least two non-
CC   contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides complementary
CC   to a portion of the HBV RNA, may be used for inhibiting HBV replication
CC   in a cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 9 G; 2 T; 4 U; 0 Other;

  Query Match           100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity  80.0%;  Pred. No. 9.9;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||:||||||||||:||::
Db          1 GACAUGAACAAGAGAUGAUU 20


RESULT 7
ADC64742/c
ID   ADC64742 standard; RNA; 39 BP.
XX
AC   ADC64742;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Hepatitis B virus DNA polymerase related RNA oligonucleotide.
XX
KW   screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
XX
OS   Synthetic.
OS   Hepatitis B virus.
XX
PN   KR2002007891-A.
XX
PD   29-JAN-2002.
XX
PF   19-JUL-2000; 2000KR-00041420.
XX
PR   19-JUL-2000; 2000KR-00041420.
XX
PA   (MOGA-) MOGAM BIOTECHNOLOGY INST.
PA   (VIRO-) VIROGEN CO LTD.
XX
PI   Ji HJ,  Jung SI,  Kim YC,  Min MG,  Ryu WS,  Yoon GS;
XX
DR   WPI; 2003-309015/30.
XX
PT   Screening of antiviral agents by protein-priming activity of hepatitis B
PT   virus DNA polymerase.
XX
PS   Disclosure; Page 12; 13pp; Korean.
XX
CC   The present invention describes a method of screening for an antiviral
CC   agent by the protein-priming activity of hepatitis B virus (HBV) DNA
CC   polymerase. Also described is developing an antiviral agent with a high
CC   selectivity to HBV which can be used for high-throughput screening. The
CC   present sequence represents an RNA oligonucleotide which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 39 BP; 5 A; 13 C; 3 G; 0 T; 18 U; 0 Other;

  Query Match           100.0%;  Score 20;  DB 10;  Length 39;
  Best Local Similarity  100.0%;  Pred. No. 10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         37 GACATGAACAAGAGATGATT 18


RESULT 8
AAD09094/c
ID   AAD09094 standard; DNA; 87 BP.
XX
AC   AAD09094;
XX
DT   04-SEP-2001  (first entry)
```

```
XX
DE   Hepatitis B virus FR1 strain genotype G HBeAg DNA fragment.
XX
KW   HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
KW   preS2; surface antigen; HBsAg; HBX protein; vaccine; HBeAg;
KW   liver disease; hepatitis; liver cancer; HBcAg; core antigen; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200138498-A2.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US032108.
XX
PR   24-NOV-1999;   99US-0167206P.
XX
PA   (PHAR-) PHARMASSET INC.
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Schinazi R,  De Gendt S,  Van Geyt C,  Zoulim F,  Fried M;
PI   Rossau R;
XX
DR   WPI; 2001-367676/38.
XX
PT   Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT   polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT   treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
PS   Claim 6; Page 57; 84pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) strain FR1,
CC   genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
CC   PreS2 and surface antigen HBsAg) and HBX proteins. HBV genotype G nucleic
CC   acids and polypeptides are useful for diagnosing, prognosing and treating
CC   infections caused by HBV genotype G. They can be used in a vaccine to
CC   treat or prevent HBV genotype G infection. The HBV genotype G derived
CC   nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC   sample or diagnosis of HBV genotype G infection. The presence of HBV
CC   genotype G statistically correlates with the presence of liver damage
CC   and/or liver cancer in the subject. The HBV genotype G core insert
CC   peptide encoding nucleic acid is useful for designing monitoring assays
CC   to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC   and HBeAg (genotype G e antigen) in patients infected with HBV. The
CC   antibodies or antigens of HBV genotype G are useful for identifying a
CC   stage of liver disease caused by HBV genotype G. The present sequence is
CC   a hepatitis B virus (HBV) strain FR1, genotype G DNA fragment encoding e
CC   antigen (HBeAg)
XX
SQ   Sequence 87 BP; 14 A; 24 C; 17 G; 32 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 4;  Length 87;
  Best Local Similarity  100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         43 GACATGAACAAGAGATGATT 24


RESULT 9
AAD09093/c
ID   AAD09093 standard; DNA; 129 BP.
XX
AC   AAD09093;
XX
DT   04-SEP-2001  (first entry)
XX
DE   Hepatitis B virus FR1 strain genotype G DNA fragment #1.
XX
KW   HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
KW   preS2; surface antigen; HBsAg; HBX protein; vaccine; liver disease;
KW   hepatitis; liver cancer; HBcAg; core antigen; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200138498-A2.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US032108.
XX
PR   24-NOV-1999;   99US-0167206P.
XX
PA   (PHAR-) PHARMASSET INC.
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Schinazi R,  De Gendt S,  Van Geyt C,  Zoulim F,  Fried M;
PI   Rossau R;
XX
DR   WPI; 2001-367676/38.
XX
PT   Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT   polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT   treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
PS   Claim 5; Page 57; 84pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) strain FR1,
CC   genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
CC   PreS2 and surface antigen HBsAg) and HBX proteins. HBV genotype G nucleic
CC   acids and polypeptides are useful for diagnosing, prognosing and treating
CC   infections caused by HBV genotype G. They can be used in a vaccine to
CC   treat or prevent HBV genotype G infection. The HBV genotype G derived
CC   nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC   sample or diagnosis of HBV genotype G infection. The presence of HBV
CC   genotype G statistically correlates with the presence of liver damage
CC   and/or liver cancer in the subject. The HBV genotype G core insert
CC   peptide encoding nucleic acid is useful for designing monitoring assays
CC   to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC   and HBeAg (genotype G e antigen) in patients infected with HBV. The
CC   antibodies or antigens of HBV genotype G are useful for identifying a
CC   stage of liver disease caused by HBV genotype G. The present sequence is
CC   a hepatitis B virus (HBV) strain FR1, genotype G DNA fragment
XX
SQ   Sequence 129 BP; 25 A; 32 C; 26 G; 46 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 4;  Length 129;
  Best Local Similarity  100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         43 GACATGAACAAGAGATGATT 24


RESULT 10
AAD27422/c
ID   AAD27422 standard; DNA; 639 BP.
XX
AC   AAD27422;
XX
DT   18-APR-2002  (first entry)
XX
DE   Hepatitis B virus (HBV) core antigen (HBcAg) encoding DNA #1.
XX
KW   Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
KW   prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KW   hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW   tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KW   dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KW   food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
KW   antiprotozoal; ds.
XX
OS   Hepatitis B virus.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag=  a
FT                   /product= "HBcAg"
XX
PN   WO200198333-A2.
XX
PD   27-DEC-2001.
XX
PF   22-JUN-2001; 2001WO-GB002817.
XX
PR   22-JUN-2000; 2000GB-00015308.
PR   06-OCT-2000; 2000GB-00024544.
XX
PA   (CELL-) CELLTECH PHARM LTD.
XX
PI   Page M,  Li J,  Pumpens P;
XX
DR   WPI; 2002-098223/13.
DR   P-PSDB; AAE17018.
XX
PT   New proteins comprising a modified hepatitis B core antigen, useful as a
PT   vaccine in prophylactic or therapeutic vaccination of the human or animal
PT   body, particularly against hepatitis B virus infection.
XX
PS   Disclosure; Page 38-39; 40pp; English.
XX
CC   The invention relates to modified proteins comprising hepatitis B virus
CC   (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC   repeats has been deleted and the protein comprising the C-terminal
CC   cysteine of HBcAg. The deleted region may be replaced by an epitope from
CC   a protein other than HBcAg, in which case the HBcAg acts as a carrier to
CC   present the epitope to the immune system. This chimeric protein or its
CC   nucleic acid is useful as a vaccine or in a method of prophylactic or
CC   therapeutic vaccination of the human or animal body, particularly against
CC   HBV. The nucleic acid encoding the protein may be used in gene therapy or
CC   DNA vaccination protocols. The chimeric protein or its nucleic acid may
CC   also be used as the basis of a prophylactic vaccine against a range of
CC   diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
CC   influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
CC   immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
CC   tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
CC   diarrhoea, meningitis or gonorrhea. The present sequence is a DNA
CC   encoding Hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 6;  Length 639;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         43 GACATGAACAAGAGATGATT 24


RESULT 11
AAD31509/c
ID   AAD31509 standard; DNA; 639 BP.
XX
AC   AAD31509;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Hepatitis B virus core antigen (HBcAg) encoding DNA.
XX
KW   Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
KW   therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
KW   polio; herpes; rabies; AIDS; foot-and-mouth disease; ds.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag=  a
FT                   /product= "HBc protein"
FT   sig_peptide     1. .87
FT                   /*tag=  b
FT   mat_peptide     88. .636
FT                   /*tag=  c
FT                   /product= "Mature HBc protein"
XX
PN   WO200177158-A1.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-GB001607.
XX
PR   07-APR-2000; 2000EP-00107118.
XX
PA   (MEDE-) MEDEVA EURO LTD.
XX
PI   Gehin A,  Gilbert R,  Stuart D,  Rowlands D;
XX
DR   WPI; 2002-239995/29.
DR   P-PSDB; AAE19793.
XX
PT   Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for the
PT   prophylactic or therapeutic treatment of humans or animals against e.g.
PT   HB virus, viral hepatitis, hepatitis C virus, influenza, or foot-and-
PT   mouth disease.
XX
PS   Disclosure; Page 23-24; 27pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) core antigen
CC   (HBcAg) fusion proteins and polynucleotides encoding such proteins.
CC   Sequences of the invention are useful in methods of prophylactic or
CC   therapeutic vaccination or to manufacture medicaments for prophylactic or
CC   therapeutic vaccination of the human or animal body against HBV, e.g.
CC   against viral hepatitis. They are also useful as a prophylactic vaccine
CC   against e.g. hepatitis c virus, influenza, polio, herpes, rabies,
CC   acquired immune deficiency syndrome (AIDS) or foot-and-mouth disease. The
CC   present sequence is a DNA encoding hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 6;  Length 639;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         43 GACATGAACAAGAGATGATT 24


RESULT 12
AAH77569/c
ID   AAH77569 standard; DNA; 655 BP.
XX
AC   AAH77569;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV genotype G strain US1 preCore/Core DNA.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW   HBeAg; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP011526.
```

XX
PR 03-DEC-1999; 99EP-00870252.
PR 07-DEC-1999; 99US-0169287P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.
XX
PS Claim 3; Fig 7; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC This genotype was found with a high prevalence in patients chronically
CC infected with HBV and residing in Europe and the USA. The invention
CC relates to a fully defined sequence of 3248 nucleotides as given in
CC specification, a sequence with 92% identity to the given sequence, or
CC sequence that is degenerate to the mentioned sequences. These
CC polynucleotides are useful for HBV genotyping. The proteins encoded by
CC the polynucleotides are useful for detecting antibodies in a biological
CC sample. Ligands that bind to the proteins and antibodies directed against
CC the proteins are useful for detecting the proteins and for detecting
CC HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC preparing a vaccine or medicament for treating HBV infections. The
CC present sequence is provided in an alignment of preCore/Core sequences of
CC an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC US6, US7, US9, US10) of HBV genotype G
XX
SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 20; DB 4; Length 655;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATGAACAAGAGATGATT 20
||||||||||||||||||||
Db 43 GACATGAACAAGAGATGATT 24

RESULT 13
AAH77568/c
ID AAH77568 standard; DNA; 655 BP.
XX
AC AAH77568;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype G strain FR2 preCore/Core DNA.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW HBeAg; ds.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP011526.
XX
PR 03-DEC-1999; 99EP-00870252.
PR 07-DEC-1999; 99US-0169287P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.
XX
PS Claim 3; Fig 7; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC This genotype was found with a high prevalence in patients chronically
CC infected with HBV and residing in Europe and the USA. The invention
CC relates to a fully defined sequence of 3248 nucleotides as given in
CC specification, a sequence with 92% identity to the given sequence, or
CC sequence that is degenerate to the mentioned sequences. These
CC polynucleotides are useful for HBV genotyping. The proteins encoded by
CC the polynucleotides are useful for detecting antibodies in a biological
CC sample. Ligands that bind to the proteins and antibodies directed against
CC the proteins are useful for detecting the proteins and for detecting
CC HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC preparing a vaccine or medicament for treating HBV infections. The
CC present sequence is provided in an alignment of preCore/Core sequences of
CC an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC US6, US7, US9, US10) of HBV genotype G
XX
SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 20; DB 4; Length 655;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATGAACAAGAGATGATT 20
||||||||||||||||||||
Db 43 GACATGAACAAGAGATGATT 24

RESULT 14
AAH77574/c
ID AAH77574 standard; DNA; 655 BP.
XX
AC AAH77574;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype G strain US10 preCore/Core DNA.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW HBeAg; ds.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP011526.
XX
PR 03-DEC-1999; 99EP-00870252.
PR 07-DEC-1999; 99US-0169287P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.

```
XX
PS   Claim 3; Fig 7; 94pp; English.
XX
CC   The invention relates to the complete nucleic acid sequence of a new
CC   human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC   This genotype was found with a high prevalence in patients chronically
CC   infected with HBV and residing in Europe and the USA. The invention
CC   relates to a fully defined sequence of 3248 nucleotides as given in
CC   specification, a sequence with 92% identity to the given sequence, or
CC   sequence that is degenerate to the mentioned sequences. These
CC   polynucleotides are useful for HBV genotyping. The proteins encoded by
CC   the polynucleotides are useful for detecting antibodies in a biological
CC   sample. Ligands that bind to the proteins and antibodies directed against
CC   the proteins are useful for detecting the proteins and for detecting
CC   HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC   preparing a vaccine or medicament for treating HBV infections. The
CC   present sequence is provided in an alignment of preCore/Core sequences of
CC   an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC   US6, US7, US9, US10) of HBV genotype G
XX
SQ   Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

  Query Match             100.0%;  Score 20;  DB 4;  Length 655;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         43 GACATGAACAAGAGATGATT 24


RESULT 15
AAH77573/c
ID   AAH77573 standard; DNA; 655 BP.
XX
AC   AAH77573;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV genotype G strain US7 preCore/Core DNA.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW   HBeAg; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP011526.
XX
PR   03-DEC-1999;   99EP-00870252.
PR   07-DEC-1999;   99US-0169287P.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Van Geyt C,  De Gendt S;
XX
DR   WPI; 2001-374785/39.
XX
PT   Novel isolated and/or purified hepatitis B virus polypeptide and
PT   polynucleotide sequences that are phylogenetically different from HBV
PT   genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT   therapy.
XX
PS   Claim 3; Fig 7; 94pp; English.
XX
CC   The invention relates to the complete nucleic acid sequence of a new
CC   human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC   This genotype was found with a high prevalence in patients chronically
CC   infected with HBV and residing in Europe and the USA. The invention
CC   relates to a fully defined sequence of 3248 nucleotides as given in
CC   specification, a sequence with 92% identity to the given sequence, or
CC   sequence that is degenerate to the mentioned sequences. These
CC   polynucleotides are useful for HBV genotyping. The proteins encoded by
CC   the polynucleotides are useful for detecting antibodies in a biological
CC   sample. Ligands that bind to the proteins and antibodies directed against
CC   the proteins are useful for detecting the proteins and for detecting
CC   HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC   preparing a vaccine or medicament for treating HBV infections. The
CC   present sequence is provided in an alignment of preCore/Core sequences of
CC   an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC   US6, US7, US9, US10) of HBV genotype G
XX
SQ   Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

  Query Match             100.0%;  Score 20;  DB 4;  Length 655;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              ||||||||||||||||||||
Db         43 GACATGAACAAGAGATGATT 24


Search completed: March 17, 2005, 06:48:41
Job time : 171.333 secs
```

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-60
Perfect score: 20
Sequence: 1 gacatgaacaagagatgatt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	18.4	92.0	763	9	CG381974	CG381974 OG1BK24TH
c 2	18.4	92.0	773	9	CG381984	CG381984 OG1BK24TV
c 3	18.4	92.0	846	9	CG373212	CG373212 OG1CZ28TH
4	18.4	92.0	892	9	CG373225	CG373225 OG1CZ28TV
c 5	17.4	87.0	280	9	BX288914	BX288914 Arabidops
6	17.4	87.0	305	6	CA520799	CA520799 KS11017B1
c 7	17.4	87.0	313	7	CF906991	CF906991 A0504F09-
8	17.4	87.0	423	5	BX837806	BX837806 BX837806
c 9	17.4	87.0	426	8	AQ183243	AQ183243 HS_3140_B
10	17.4	87.0	442	7	CN958899	CN958899 6399_1001
11	17.4	87.0	450	8	AZ654527	AZ654527 1M0528D16
12	17.4	87.0	509	3	AK015232	AK015232 Mus muscu
c 13	17.4	87.0	527	8	AQ961278	AQ961278 LERFK20TR
14	17.4	87.0	570	7	CF198522	CF198522 EST0117 T
c 15	17.4	87.0	574	7	CF909462	CF909462 A0536H03-
16	17.4	87.0	616	7	CK517114	CK517114 rswjb0_00
c 17	17.4	87.0	628	8	BH366000	BH366000 CH230-112
18	17.4	87.0	631	2	BE388774	BE388774 601283896
c 19	17.4	87.0	652	8	BZ898091	BZ898091 CH240_13N
20	17.4	87.0	682	9	CE163565	CE163565 tigr-gss-
c 21	17.4	87.0	683	7	CO817689	CO817689 FA_SEa001
22	17.4	87.0	721	8	AQ961277	AQ961277 LERFK20TF
c 23	17.4	87.0	794	9	CL809904	CL809904 OR_CBa002
24	17.4	87.0	821	2	BF678287	BF678287 602084906
c 25	17.4	87.0	859	8	CC090167	CC090167 CSU-K33r.
26	17.4	87.0	870	8	CC131380	CC131380 NDL.76K8.
c 27	17.4	87.0	885	2	BF541940	BF541940 602068444
c 28	17.4	87.0	903	8	CC068329	CC068329 CSU-K33r.
29	17.4	87.0	1340	3	CNS0A5V5	BX823129 Arabidops
30	17	85.0	472	2	AW760013	AW760013 sl56h09.y
c 31	17	85.0	532	1	AL819446	AL819446 AL819446
c 32	17	85.0	577	4	BI378081	BI378081 BFLG3_001
33	17	85.0	687	9	AG140746	AG140746 Pan trogl
34	17	85.0	730	5	BX114353	BX114353 BX114353
c 35	17	85.0	768	9	CC786410	CC786410 ZMMBBb015
36	17	85.0	778	7	CO368799	CO368799 RTK1_42_H
37	17	85.0	808	9	CL543431	CL543431 OB__Ba006
38	17	85.0	1013	9	CNS06RAQ	AL411720 T3 end of
39	16.8	84.0	124	4	BI127956	BI127956 G068P81Y
40	16.8	84.0	170	4	BI128183	BI128183 G072P24Y
41	16.8	84.0	195	4	BG125265	BG125265 EST470911
42	16.8	84.0	195	4	BG733602	BG733602 Cs-20d Ci
c 43	16.8	84.0	218	2	BE428564	BE428564 MTD008.D0
c 44	16.8	84.0	234	4	BI473621	BI473621 fp39h03.y
45	16.8	84.0	251	8	BZ385056	BZ385056 SALK_1363

ALIGNMENTS

RESULT 1
CG381974
LOCUS CG381974 763 bp DNA linear GSS 26-AUG-2003
DEFINITION OG1BK24TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0724C23, genomic survey sequence.
ACCESSION CG381974
VERSION CG381974.1 GI:34299241
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 763)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OG1BK24TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .763
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0724C23"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 763;
Best Local Similarity 95.0%; Pred. No. 4.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GACATGAACAAGAGATGATT 20
              | ||||||||||||||||||
Db        572 GGCATGAACAAGAGATGATT 591
```

```
RESULT 2
CG381984/c
LOCUS       CG381984                773 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OG1BK24TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0724C23,
            genomic survey sequence.
ACCESSION   CG381984
VERSION     CG381984.1  GI:34299251
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1BK24TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .773
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0724C23"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 9;  Length 773;
  Best Local Similarity   95.0%;  Pred. No. 4.3e+02;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              | ||||||||||||||||||
Db        668 GGCATGAACAAGAGATGATT 649


RESULT 3
CG373212/c
LOCUS       CG373212                846 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OG1CZ28TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0734E08,
            genomic survey sequence.
ACCESSION   CG373212
VERSION     CG373212.1  GI:34290479
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1CZ28TV
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0734E08"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 9;  Length 846;
  Best Local Similarity   95.0%;  Pred. No. 4.3e+02;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              | ||||||||||||||||||
Db        540 GGCATGAACAAGAGATGATT 521


RESULT 4
CG373225
LOCUS       CG373225                892 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OG1CZ28TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0734E08,
            genomic survey sequence.
ACCESSION   CG373225
VERSION     CG373225.1  GI:34290492
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1CZ28TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .892
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0734E08"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 9;  Length 892;
  Best Local Similarity   95.0%;  Pred. No. 4.4e+02;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGATT 20
              | ||||||||||||||||||
```

```
Db        353 GGCATGAACAAGAGATGATT 372


RESULT 5
BX288914/c
LOCUS       BX288914                280 bp    DNA     linear   GSS 02-APR-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-421D09-018141,
            genomic survey sequence.
ACCESSION   BX288914
VERSION     BX288914.1  GI:28887910
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
  TITLE     GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
  JOURNAL   Bioinformatics 19 (11), 1441-1442 (2003)
  MEDLINE   22755829
   PUBMED   12874060
REFERENCE   2
  AUTHORS   Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
            Weisshaar,B.
  TITLE     An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
            flanking sequence tag-based reverse genetics
  JOURNAL   Plant Mol. Biol. 53 (1-2), 247-259 (2003)
  MEDLINE   23117147
   PUBMED   14756321
REFERENCE   3
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
            Weisshaar,B.
  TITLE     High-throughput generation of sequence indexes from T-DNA
            mutagenized Arabidopsis thaliana lines
  JOURNAL   BioTechniques 35 (6), 1164-1168 (2003)
   PUBMED   14682050
REFERENCE   4  (bases 1 to 280)
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence has been recovered from the left border of the T-DNA.
            It indicates an insertion within the locus defined by BAC clone
            f3k23. Details on the protocols used for generation of the sequence
            are described in References 1-3. The sequences are generated at the
            MPI for Plant Breeding Research in the context of the GABI-Kat
            project. GABI-Kat is part of the German Plant Genomics program
            designated 'GABI'. Information on line availability can be found
            at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .280
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-421D09-018141"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 280;
  Best Local Similarity   94.7%;  Pred. No. 1.1e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 ACATGAACAAGAGATGATT 20
              ||||||||||||||| ||||
Db         80 ACATGAACAAGAGAAGATT 62


RESULT 6
CA520799
LOCUS       CA520799                305 bp    mRNA    linear   EST 15-NOV-2002
DEFINITION  KS11017B12 KS11 Capsicum annuum cDNA, mRNA sequence.
ACCESSION   CA520799
VERSION     CA520799.1  GI:25034824
KEYWORDS    EST.
SOURCE      Capsicum annuum
  ORGANISM  Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Capsicum.
REFERENCE   1  (bases 1 to 305)
  AUTHORS   Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S.,
            Hur,C.-G. and Choi,D.
  TITLE     Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
            annuum L.) and Sequence Analysis in Relation to Hypersensitive
            Response Against Pathogen
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Doil Choi
            Genome Research Center and National Center for Genome Information
            Korea Research Institute of Bioscience and Biotechnology
            P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea
            Tel: 82-42-860-4340
            Fax: 82-42-860-4309
            Email: doil@mail.kribb.re.kr
            Plate: 017  row: B  column: 12.
FEATURES             Location/Qualifiers
     source          1. .305
                     /organism="Capsicum annuum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4072"
                     /clone_lib="KS11"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 6;  Length 305;
  Best Local Similarity   94.7%;  Pred. No. 1.1e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGAT 19
              ||||||||||||||| ||
Db        212 GACATGAACAAGAGATCAT 230


RESULT 7
CF906991/c
LOCUS       CF906991                313 bp    mRNA    linear   EST 05-NOV-2003
DEFINITION  A0504F09-5 NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)
            Mus musculus cDNA clone NIA:A0504F09 IMAGE:30743204 5', mRNA
            sequence.
ACCESSION   CF906991
VERSION     CF906991.1  GI:38177928
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
  TITLE     Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
  JOURNAL   Genome Res. 11 (9), 1553-1558 (2001)
  MEDLINE   21429098
   PUBMED   11544199
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
```

```
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: A0504  row: F  column: 09
            Seq primer: M13 Reverse
            High quality sequence stop: 313
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .313
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C3H/He mice"
                     /db_xref="niaEST:A0504F09-5"
                     /db_xref="taxon:10090"
                     /clone="NIA:A0504F09 IMAGE:30743204"
                     /dev_stage="9-15C cells"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse Mesenchymal Stem Cell cDNA Library
                     (Long 1)"
                     /note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
                     Site_2: NotI; Mouse cDNA project by the Laboratory of
                     Genetics, National Institute on Aging (NIA), Intramural
                     Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
                     This is a long-transcript enriched cDNA library (Ref.
                     Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
                     RNAs were obtained from Dr. Akihiro Umezawa (Keio
                     University School of Medicine, Japan). Double-stranded
                     cDNAs were synthesized with an Oligo(dT) primer
                     [Invitrogen:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from
                     2.2 ug of total RNA, treated with T4 DNA polymerase, and
                     purified by ethanol-precipitation. The cDNAs were ligated
                     to Lone-linker LL-Sal4, purified by phenol/chloroform, and
                     separated from free linkers by Centricon 100. Then, the
                     cDNAs were amplified by long-range high fidelity PCR using
                     Ex Taq polymerase (Takara) with a primer Sal4-S. The
                     products were purified by phenol/chloroform and Centricon
                     100. The cDNAs were digested with SalI and NotI enzymes
                     and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
                     vector. The DH10B E. coli host was transformed with the
                     ligation mixture by the standard chemical method. The
                     average insert size is about 2.5 kb. The library was
                     constructed by Yulan Piao."
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 7;  Length 313;
  Best Local Similarity   94.7%;  Pred. No. 1.1e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGAT 19
              || |||||||||||||||||
Db        289 GAAATGAACAAGAGATGAT 271


RESULT 8
BX837806
LOCUS       BX837806               423 bp    mRNA    linear   EST 11-FEB-2004
DEFINITION  BX837806 Arabidopsis thaliana Hormone Treated Callus Col-0
            Arabidopsis thaliana cDNA clone GSLTPGH21ZC04 5PRIM, mRNA sequence.
ACCESSION   BX837806
VERSION     BX837806.1  GI:42531889
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
            Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
            Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
  TITLE     Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
            A Combined Approach to Evaluate and Improve Arabidopsis Genome
            Annotation
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            The sequences are based on single pass reads.
            Life Technologies (a division of Invitrogen) members carried out
            full-length librairies construction : Temple G.
            Genoscope members carried out sequencing and annotation : Castelli
            V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
            Schachter V., Weissenbach J., Salanoubat M.
            URGV INRA : Clepet C., Caboche M.
            Annotation is based on the June 2003 version of the Arabidopsis
            genome released by MIPS (Munich Information center for Protein
            Sequences).
            http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
            http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Col-0"
                     /db_xref="taxon:3702"
                     /clone="GSLTPGH21ZC04"
                     /tissue_type="Hormone Treated Callus"
                     /clone_lib="Arabidopsis thaliana Hormone Treated Callus
                     Col-0"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 5;  Length 423;
  Best Local Similarity   94.7%;  Pred. No. 1.2e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 ACATGAACAAGAGATGATT 20
              ||||||| ||||||||||| |
Db        403 ACATGAAGAAGAGATGATT 421


RESULT 9
AQ183243/c
LOCUS       AQ183243               426 bp    DNA     linear   GSS 01-NOV-1998
DEFINITION  HS_3140_B2_B02_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3140 Col=4 Row=D, genomic survey
            sequence.
ACCESSION   AQ183243
VERSION     AQ183243.1  GI:3580610
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3140  row: D  column: 4
            Class: BAC ends
            High quality sequence stop: 426.
FEATURES             Location/Qualifiers
```

```
     source          1. .426
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3140 Col=4 Row=D"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 8;  Length 426;
  Best Local Similarity   94.7%;  Pred. No. 1.2e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 ACATGAACAAGAGATGATT 20
              |||||||||||||||||| |
Db        273 ACATGAACAAGAGATGACT 255


RESULT 10
CN958899
LOCUS       CN958899                442 bp    mRNA    linear   EST 08-JUN-2004
DEFINITION  6399_100122_44 Fundulus Heteroclitus Liver Fundulus heteroclitus
            cDNA, mRNA sequence.
ACCESSION   CN958899
VERSION     CN958899.1  GI:48440488
KEYWORDS    EST.
SOURCE      Fundulus heteroclitus (killifish)
  ORGANISM  Fundulus heteroclitus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Cyprinodontiformes; Fundulidae; Fundulus.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   Crawford,D.L., Oleksiak,M.F., Kolell,K.J., Paschall,J., VanWye,J.,
            Roach,J.L. and Whitehead,J.A.
  TITLE     Fundulus Functional Genomics: EST Database for Teleost Fish
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Crawford, Douglas L.
            Marine Genomics - Crawford Lab
            Rosenstiel School of Marine and Atmospheric Science - University of
            Miami
            4600 Rickenbacker Causeway, Miami, FL 33149-1098 USA
            Tel: 305 361 4121
            Email: dcrawford@rsmas.Miami.edu
            Database Web Interface
            http://genomics.rsmas.miami.edu/funnybase/super_craw3/
            Plate: 100122  row: E  column: 6.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Fundulus heteroclitus"
                     /mol_type="mRNA"
                     /db_xref="taxon:8078"
                     /tissue_type="Liver"
                     /clone_lib="Fundulus Heteroclitus Liver"
                     /note="Organ: Liver"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 7;  Length 442;
  Best Local Similarity   94.7%;  Pred. No. 1.2e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGAT 19
              | ||||||||||||||||||
Db        201 GTCATGAACAAGAGATGAT 219


RESULT 11
AZ654527
LOCUS       AZ654527                450 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0528D16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0528D16 R, genomic survey sequence.
ACCESSION   AZ654527
VERSION     AZ654527.1  GI:11791673
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 450)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0528  row: D  column: 16
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .450
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0528D16"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 8;  Length 450;
  Best Local Similarity   94.7%;  Pred. No. 1.2e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACATGAACAAGAGATGAT 19
              || ||||||||||||||||
Db        132 GAAATGAACAAGAGATGAT 150


RESULT 12
AK015232
LOCUS       AK015232                509 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4930429C20 product:unclassifiable, full insert
```

sequence.
ACCESSION AK015232
VERSION AK015232.1 GI:12853489
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 509)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
FEATURES Location/Qualifiers
source 1. .509
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4930429C20"
/db_xref="taxon:10090"
/clone="4930429C20"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature 1. .509
/note="unclassifiable"
ORIGIN

Query Match 87.0%; Score 17.4; DB 3; Length 509;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ACATGAACAAGAGATGATT 20
| ||||||||||||||||||
Db 443 AGATGAACAAGAGATGATT 461

RESULT 13
AQ961278/c
LOCUS AQ961278 527 bp DNA linear GSS 28-JAN-2000
DEFINITION LERFK20TR LERA Arabidopsis thaliana genomic clone LERFK20, genomic survey sequence.
ACCESSION AQ961278
VERSION AQ961278.1 GI:6788979
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 527)
AUTHORS Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T., Feldblyum,T., Liang,F., Creasy,T. and Fraser,C.M.
TITLE Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
JOURNAL Unpublished (2000)
COMMENT Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: at@tigr.org
For addtional information, see http://www.tigr.org/tdb/at/at.html
Seq primer: TR
Class: shotgun.
FEATURES Location/Qualifiers
source 1. .527
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LERFK20"
/clone_lib="LERA"
/note="Organ: Leaf; Vector: pHOS1; Total genomic DNA was

sheared to 0.9-1 Kbp before ligation."
ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 527;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 ACATGAACAAGAGATGATT 20
              ||||||||||||||| ||||
Db        160 ACATGAACAAGAGAAGATT 142
```

RESULT 14
CF198522
LOCUS CF198522 570 bp mRNA linear EST 01-AUG-2003
DEFINITION EST0117 Tamarix androssowii leaf Tamarix androssowii cDNA, mRNA sequence.
ACCESSION CF198522
VERSION CF198522.1 GI:33392895
KEYWORDS EST.
SOURCE Tamarix androssowii
ORGANISM Tamarix androssowii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Tamaricaceae; Tamarix.
REFERENCE 1 (bases 1 to 570)
AUTHORS Wang,Y., Yang,C., Jiang,J., Liu,G., Wu,J. and Liu,Z.
TITLE EST aquired from cDNA library of Tamarix androssowii treated with NaHCO3
JOURNAL Unpublished (2003)
COMMENT Contact: Yucheng Wang
Forestry Source and Environment College
Northeast Forestry University
Hexing 26, Harbin, Heilongjiang, 150040, P.R. China
Tel: 086-451-2190607
Email: WANGYUCHENG1029@YAHOO.COM.CN.
FEATURES Location/Qualifiers
source 1. .570
/organism="Tamarix androssowii"
/mol_type="mRNA"
/db_xref="taxon:189785"
/tissue_type="leaf"
/clone_lib="Tamarix androssowii leaf"
ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 570;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 ACATGAACAAGAGATGATT 20
              ||||||||||||| |||||
Db        143 ACATGAACAAGAGGTGATT 161
```

RESULT 15
CF909462/c
LOCUS CF909462 574 bp mRNA linear EST 05-NOV-2003
DEFINITION A0536H03-5 NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1) Mus musculus cDNA clone NIA:A0536H03 IMAGE:30746294 5', mRNA sequence.
ACCESSION CF909462
VERSION CF909462.1 GI:38180399
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 574)
AUTHORS Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
TITLE Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method
JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098
PUBMED 11544199
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: A0536 row: H column: 03
Seq primer: M13 Reverse
High quality sequence stop: 574
POLYA=No.
FEATURES Location/Qualifiers
source 1. .574
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H/He mice"
/db_xref="niaEST:A0536H03-5"
/db_xref="taxon:10090"
/clone="NIA:A0536H03 IMAGE:30746294"
/dev_stage="9-15C cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Akihiro Umezawa (Keio University School of Medicine, Japan). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3'] from 2.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."
ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 574;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GACATGAACAAGAGATGAT 19
              || |||||||||||||||
Db        510 GAAATGAACAAGAGATGAT 492
```

Search completed: March 17, 2005, 11:07:46
Job time : 1390.27 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-61
Perfect score: 20
Sequence: 1 gacaugaacaagagaugauu 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	30	6	AR027810	AR027810 Sequence
2	20	100.0	30	6	AR027841	AR027841 Sequence
c 3	20	100.0	87	6	AX151115	AX151115 Sequence
c 4	20	100.0	99	14	HPBPRECA	M76687 Hepatitis B
c 5	20	100.0	99	14	HPBPRECB	M76688 Hepatitis B
c 6	20	100.0	99	14	HPBPRECC	M76689 Hepatitis B
c 7	20	100.0	99	14	HPBPRECD	M76690 Hepatitis B
c 8	20	100.0	99	14	HPBPRECE	M76691 Hepatitis B
c 9	20	100.0	99	14	HPBPRECF	M76692 Hepatitis B
c 10	20	100.0	99	14	HPBPRECG	M76693 Hepatitis B
c 11	20	100.0	99	14	HPBPRECH	M76694 Hepatitis B
c 12	20	100.0	99	14	HPBPRECI	M76695 Hepatitis B
c 13	20	100.0	99	14	HPBPRECM	M76699 Hepatitis B
c 14	20	100.0	129	6	AX151114	AX151114 Sequence
c 15	20	100.0	150	14	AF528205	AF528205 Hepatitis
c 16	20	100.0	150	14	AF528206	AF528206 Hepatitis
c 17	20	100.0	150	14	AF528207	AF528207 Hepatitis
c 18	20	100.0	150	14	AF528208	AF528208 Hepatitis
c 19	20	100.0	150	14	AF528209	AF528209 Hepatitis
c 20	20	100.0	150	14	AF528210	AF528210 Hepatitis
c 21	20	100.0	150	14	AF528211	AF528211 Hepatitis
c 22	20	100.0	150	14	AF528212	AF528212 Hepatitis
c 23	20	100.0	150	14	AF528213	AF528213 Hepatitis
c 24	20	100.0	150	14	AF528214	AF528214 Hepatitis
c 25	20	100.0	150	14	AF528215	AF528215 Hepatitis
c 26	20	100.0	150	14	AF528216	AF528216 Hepatitis
c 27	20	100.0	150	14	AF528217	AF528217 Hepatitis
c 28	20	100.0	150	14	AF528218	AF528218 Hepatitis
c 29	20	100.0	150	14	AF528219	AF528219 Hepatitis
c 30	20	100.0	150	14	AF528220	AF528220 Hepatitis
c 31	20	100.0	150	14	AF528221	AF528221 Hepatitis
c 32	20	100.0	150	14	AF528222	AF528222 Hepatitis
c 33	20	100.0	150	14	AF528224	AF528224 Hepatitis
c 34	20	100.0	150	14	AF528225	AF528225 Hepatitis
c 35	20	100.0	150	14	AF528226	AF528226 Hepatitis
c 36	20	100.0	150	14	AF528227	AF528227 Hepatitis
c 37	20	100.0	150	14	AF528228	AF528228 Hepatitis
c 38	20	100.0	150	14	AF528229	AF528229 Hepatitis
c 39	20	100.0	150	14	AF528231	AF528231 Hepatitis
c 40	20	100.0	150	14	AF528232	AF528232 Hepatitis
c 41	20	100.0	150	14	AF528233	AF528233 Hepatitis
c 42	20	100.0	150	14	AF528234	AF528234 Hepatitis
c 43	20	100.0	150	14	AF528235	AF528235 Hepatitis
c 44	20	100.0	150	14	AF528236	AF528236 Hepatitis
c 45	20	100.0	150	14	AF528237	AF528237 Hepatitis

ALIGNMENTS

RESULT 1
AR027810
LOCUS AR027810 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 8 from patent US 5856459.
ACCESSION AR027810
VERSION AR027810.1 GI:5938630
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and Mills,J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 8 05-JAN-1999;
FEATURES Location/Qualifiers
source 1. .30
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAGAGAUGAUU 20
||||:||||||||||::
Db 1 GACATGAACAAGAGATGATT 20

RESULT 2
AR027841
LOCUS AR027841 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 39 from patent US 5856459.
ACCESSION AR027841
VERSION AR027841.1 GI:5938661
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and

```
            Mills,J.S.
  TITLE     Oligonucleotides specific for hepatitis B virus
  JOURNAL   Patent: US 5856459-A 39 05-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 56;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db          1 GACATGAACAAGAGATGATT 20


RESULT 3
AX151115/c
LOCUS       AX151115                  87 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 4 from Patent WO0138498.
ACCESSION   AX151115
VERSION     AX151115.1  GI:14533317
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F.,
            Fried,M. and Rossau,R.
  TITLE     A new genotype of hepatitis b virus
  JOURNAL   Patent: WO 0138498-A 4 31-MAY-2001;
            Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 87;
  Best Local Similarity   80.0%;  Pred. No. 51;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         43 GACATGAACAAGAGATGATT 24


RESULT 4
HPBPRECA/c
LOCUS       HPBPRECA                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type1 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76687
VERSION     M76687.1  GI:485341
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45507.1"
                     /db_xref="GI:485342"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 1 (creates internal stop
                     codon)"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 5
HPBPRECB/c
LOCUS       HPBPRECB                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 2precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76688
VERSION     M76688.1  GI:485343
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     variation       2
                     /note="c in wt; t in virus type 2"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45508.1"
                     /db_xref="GI:485344"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 2 (creates internal stop
                     codon)"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 6
HPBPRECC/c
LOCUS       HPBPRECC                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 3precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76689
VERSION     M76689.1  GI:485345
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     variation       6
                     /note="c in wt; t in virus type 3"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45509.1"
                     /db_xref="GI:485346"
                     /translation="MQLFHLCLIISCSCPTFQASKLCLGWL"
     variation       58
                     /gene="C"
                     /note="g in wt; t in virus type 3 (val to phe)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 3 (creates internal stop
                     codon)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 7
HPBPRECD/c
LOCUS       HPBPRECD                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 4 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76690
VERSION     M76690.1  GI:485347
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45510.1"
                     /db_xref="GI:485348"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 4 (creates internal stop
                     codon)"
     variation       95
                     /note="g in wt; a in virus type 4 (gly to asp)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 8
HPBPRECE/c
LOCUS       HPBPRECE                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 5 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76691
VERSION     M76691.1  GI:485349
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     CDS             10. .93
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45511.1"
```

```
                     /db_xref="GI:485350"
                     /translation="MQLFHLCLIISCSCPTVQPSKLCLGWL"
     variation       64. .67
                     /gene="C"
                     /note="gcc in wt; ccg in virus type 5 (ala to pro)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 5 (creates internal stop
                     codon)"
     variation       95
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:|||||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 9
HPBPRECF/c
LOCUS       HPBPRECF                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 6 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76692
VERSION     M76692.1  GI:485351
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .99
                     /gene="C"
     misc_feature    10. .99
                     /gene="C"
                     /product="precore protein"
                     /note="putative cds"
     variation       11
                     /gene="C"
                     /note="t in wt; c in virus type 6 (loss of start codon)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:|||||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 10
HPBPRECG/c
LOCUS       HPBPRECG                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 7 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76693
VERSION     M76693.1  GI:485352
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     misc_feature    10. .93
                     /gene="C"
                     /product="precore protein"
                     /standard_name="pre-C region note: putative CDS"
     variation       10
                     /gene="C"
                     /note="a in wt; t in virus type 7 (loss of start codon)"
     variation       14
                     /gene="C"
                     /note="a in wt; g in virus type 7 (gln to arg)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 7 ( creates internal stop
                     codon)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:|||||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 11
HPBPRECH/c
LOCUS       HPBPRECH                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 8 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76694
VERSION     M76694.1  GI:485353
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     misc_feature    10. .93
```

```
                     /gene="C"
                     /product="precore protein"
                     /standard_name="pre-C region note: putative CDS"
     variation       12
                     /gene="C"
                     /note="g in wt; t in virus type 8 (loss of start codon)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 8 (creates internal stop
                     codon)"
     variation       95
                     /note="g in wt; a in virus type 8 (gly to asp)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 12
HPBPRECI/c
LOCUS       HPBPRECI                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 9 precore protein (pre-C region, C) gene, 5'
            end.
ACCESSION   M76695
VERSION     M76695.1  GI:485354
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .93
                     /gene="C"
     misc_feature    10. .93
                     /gene="C"
                     /product="precore protein"
                     /standard_name="pre-C region note: putative CDS"
     variation       13
                     /gene="C"
                     /note="c in wt; t in virus type 9 (creates internal stop
                     codon)"
     variation       92
                     /gene="C"
                     /note="g in wt; a in virus type 9 (creates internal stop
                     codon)"
     variation       95
                     /note="g in wt; a in virus type 9 (gly to asp)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 13
HPBPRECM/c
LOCUS       HPBPRECM                  99 bp    DNA     linear   VRL 11-MAY-1994
DEFINITION  Hepatitis B virus type 13 precore protein (pre-C region, C) gene,
            5' end.
ACCESSION   M76699
VERSION     M76699.1  GI:485361
KEYWORDS    e antigen; precore protein; tolerogen.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
            Will,H.
  TITLE     Prevalence and type of pre-C HBV mutants in anti-HBe positive
            carriers with chronic liver disease in a highly endemic area
  JOURNAL   Virology 183 (2), 840-844 (1991)
  MEDLINE   91306476
   PUBMED   1853582
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            10. .99
                     /gene="C"
     CDS             10. .>99
                     /gene="C"
                     /standard_name="pre-C region"
                     /codon_start=1
                     /product="precore protein"
                     /protein_id="AAA45515.1"
                     /db_xref="GI:485362"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWLWDM"
     variation       95
                     /gene="C"
                     /note="g in wt; a in virus type 13 (gly to asp)"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 50;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         52 GACATGAACAAGAGATGATT 33


RESULT 14
AX151114/c
LOCUS       AX151114                 129 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3 from Patent WO0138498.
ACCESSION   AX151114
VERSION     AX151114.1  GI:14533316
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F.,
            Fried,M. and Rossau,R.
  TITLE     A new genotype of hepatitis b virus
  JOURNAL   Patent: WO 0138498-A 3 31-MAY-2001;
            Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES             Location/Qualifiers
     source          1. .129
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN
```

```
  Query Match             100.0%;  Score 20;  DB 6;  Length 129;
  Best Local Similarity   80.0%;  Pred. No. 49;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         43 GACATGAACAAGAGATGATT 24


RESULT 15
AF528205/c
LOCUS       AF528205                150 bp    DNA     linear   VRL 31-JUL-2003
DEFINITION  Hepatitis B virus ASC1123 core antigen precursor, gene, partial
            cds.
ACCESSION   AF528205
VERSION     AF528205.1  GI:32810971
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 150)
  AUTHORS   Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
  TITLE     Comparative evaluation of HBV precore and basal core promoter
            mutants in Indian patients with diverse clinical manifestations
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150)
  AUTHORS   Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUL-2002) Hepatitis Division, National Institute of
            Virology, 20-A, Dr Ambedkar Road, Pune, Maharashtra 411001, India
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="Hepatitis B virus"
                     /proviral
                     /mol_type="genomic DNA"
                     /isolate="ASC1123"
                     /isolation_source="asymptomatic HBsAg carrier"
                     /specific_host="Homo sapiens"
                     /db_xref="taxon:10407"
                     /country="India"
     misc_feature    <1. .>150
                     /note="contains partial basal core promoter"
     CDS             64. .>150
                     /note="contains complete precore region"
                     /codon_start=1
                     /product="core antigen precursor"
                     /protein_id="AAP87556.1"
                     /db_xref="GI:32810972"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWLXG"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 150;
  Best Local Similarity   80.0%;  Pred. No. 48;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db        106 GACATGAACAAGAGATGATT 87


Search completed: March 17, 2005, 08:14:16
Job time : 683.733 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-61
Perfect score: 20
Sequence: 1 gacaugaacaagagaugauu 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	2	AAT72566	Aat72566 Hepatitis
2		20	100.0	20	2	AAT72565	Aat72565 Hepatitis
3		20	100.0	30	2	AAT72562	Aat72562 Hepatitis
4		20	100.0	30	2	AAT72563	Aat72563 Hepatitis
5		20	100.0	30	2	AAT72616	Aat72616 Hepatitis
6		20	100.0	30	2	AAT72617	Aat72617 Hepatitis
c 7		20	100.0	39	10	ADC64742	Adc64742 Hepatitis
c 8		20	100.0	87	4	AAD09094	Aad09094 Hepatitis
c 9		20	100.0	129	4	AAD09093	Aad09093 Hepatitis
c 10		20	100.0	639	6	AAD27422	Aad27422 Hepatitis
c 11		20	100.0	639	6	AAD31509	Aad31509 Hepatitis
c 12		20	100.0	655	4	AAH77569	Aah77569 HBV genot
c 13		20	100.0	655	4	AAH77568	Aah77568 HBV genot
c 14		20	100.0	655	4	AAH77574	Aah77574 HBV genot
c 15		20	100.0	655	4	AAH77573	Aah77573 HBV genot
c 16		20	100.0	655	4	AAH77570	Aah77570 HBV genot
c 17		20	100.0	655	4	AAH77571	Aah77571 HBV genot
c 18		20	100.0	664	4	AAH77572	Aah77572 HBV genot
c 19		20	100.0	669	12	ADO07220	Ado07220 Hepatitis
c 20		20	100.0	673	4	AAD09092	Aad09092 Hepatitis
c 21		20	100.0	675	4	AAH77563	Aah77563 HBV preCo
c 22		20	100.0	681	4	AAH77567	Aah77567 HBV genot
c 23		20	100.0	1395	2	AAV82688	Aav82688 Fulminant
c 24		20	100.0	1400	2	AAV82687	Aav82687 Fulminant
c 25		20	100.0	1445	2	AAV82692	Aav82692 Fulminant
c 26		20	100.0	1445	2	AAV82685	Aav82685 Fulminant
c 27		20	100.0	1445	2	AAV82690	Aav82690 Fulminant
c 28		20	100.0	1445	2	AAV82684	Aav82684 Fulminant
c 29		20	100.0	1500	2	AAV82695	Aav82695 Fulminant
c 30		20	100.0	1500	2	AAV82683	Aav82683 Fulminant
c 31		20	100.0	1500	2	AAV82694	Aav82694 Fulminant
c 32		20	100.0	1500	2	AAV82686	Aav82686 Fulminant
c 33		20	100.0	1500	2	AAV82706	Aav82706 Wild type
c 34		20	100.0	1500	2	AAV82689	Aav82689 Fulminant
c 35		20	100.0	1500	2	AAV82693	Aav82693 Fulminant
c 36		20	100.0	2342	1	AAN93072	Aan93072 Sequence
c 37		20	100.0	2743	1	AAN00003	Aan00003 Sequence
c 38		20	100.0	2743	2	AAQ04799	Aaq04799 Recombina
c 39		20	100.0	3180	4	AAH42375	Aah42375 Nucleotid
c 40		20	100.0	3182	6	AAD31765	Aad31765 Hepatitis
c 41		20	100.0	3182	9	ACA62422	Aca62422 Hepatitis
c 42		20	100.0	3182	10	AAD60866	Aad60866 Hepatitis
c 43		20	100.0	3220	3	AAZ88924	Aaz88924 Hepatitis
c 44		20	100.0	3248	4	AAD09091	Aad09091 Hepatitis
c 45		20	100.0	3248	4	AAH77562	Aah77562 HBV genot

ALIGNMENTS

RESULT 1
AAT72566
ID AAT72566 standard; RNA; 20 BP.
XX
AC AAT72566;
XX
DT 03-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV46MYb.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1. .20
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
FT modified_base 1
FT /*tag= b
FT /mod_base= gm
FT modified_base 2
FT /*tag= c
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 3
FT /*tag= d
FT /mod_base= cm
FT modified_base 4
FT /*tag= e
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 5
FT /*tag= f
FT /mod_base= um
FT modified_base 6
FT /*tag= g
FT /mod_base= gm
FT modified_base 7
FT /*tag= h
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 8

```
FT                   /*tag=  i
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   9
FT                   /*tag=  j
FT                   /mod_base= cm
FT   modified_base   10
FT                   /*tag=  k
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   11
FT                   /*tag=  l
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   12
FT                   /*tag=  m
FT                   /mod_base= gm
FT   modified_base   13
FT                   /*tag=  n
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   14
FT                   /*tag=  o
FT                   /mod_base= gm
FT   modified_base   15
FT                   /*tag=  p
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   16
FT                   /*tag=  q
FT                   /mod_base= um
FT   modified_base   17
FT                   /*tag=  r
FT                   /mod_base= gm
FT   modified_base   18
FT                   /*tag=  s
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   19
FT                   /*tag=  t
FT                   /mod_base= um
FT   modified_base   20
FT                   /*tag=  u
FT                   /mod_base= um
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV46MYb
CC   which is complementary to a portion of the hepatitis B virus (HBV) RNA.
CC   The antisense oligonucleotide may be used to detect the presence of HBV
CC   in a sample. The antisense oligonucleotide, and oligonucleotides
CC   containing a sequence which is complementary to at least two non-
CC   contiguous regions of an HBV nucleic acid, may be used for inhibiting HBV
CC   replication in a cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 9 A; 2 C; 5 G; 0 T; 4 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 9.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||||||||||||||||||
Db          1 GACAUGAACAAGAGAUGAUU 20


RESULT 2
AAT72565
ID   AAT72565 standard; DNA; 20 BP.
XX
AC   AAT72565;
XX
DT   03-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV46Yb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .20
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV46Yb which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 9 A; 2 C; 5 G; 4 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 9.6;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||:||::
Db          1 GACATGAACAAGAGATGATT 20


RESULT 3
AAT72562
```

```
ID   AAT72562 standard; DNA; 30 BP.
XX
AC   AAT72562;
XX
DT   03-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV88b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag= a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV88b which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 10 G; 5 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 9.9;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:|||||||||:||::
Db          1 GACATGAACAAGAGATGATT 20


RESULT 4
AAT72563
ID   AAT72563 standard; DNA; 30 BP.
XX
AC   AAT72563;
XX
DT   03-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV88Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag= a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        1. .20
FT                   /*tag= b
FT                   /note= "2'-OMe RNA"
FT   modified_base   1
FT                   /*tag= c
FT                   /mod_base= gm
FT   modified_base   2
FT                   /*tag= d
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   3
FT                   /*tag= e
FT                   /mod_base= cm
FT   modified_base   4
FT                   /*tag= f
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   5
FT                   /*tag= g
FT                   /mod_base= um
FT   modified_base   6
FT                   /*tag= h
FT                   /mod_base= gm
FT   modified_base   7
FT                   /*tag= i
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   8
FT                   /*tag= j
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   9
FT                   /*tag= k
FT                   /mod_base= cm
FT   modified_base   10
FT                   /*tag= l
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   11
FT                   /*tag= m
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   12
FT                   /*tag= n
FT                   /mod_base= gm
FT   modified_base   13
FT                   /*tag= o
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   14
FT                   /*tag= p
FT                   /mod_base= gm
FT   modified_base   15
FT                   /*tag= q
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   16
FT                   /*tag= r
FT                   /mod_base= um
FT   modified_base   17
FT                   /*tag= s
FT                   /mod_base= gm
FT   modified_base   18
FT                   /*tag= s
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   19
FT                   /*tag= u
FT                   /mod_base= um
FT   modified_base   20
FT                   /*tag= v
```

```
FT                   /mod_base= um
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV88Mb which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 10 G; 1 T; 4 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity  100.0%;  Pred. No. 9.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||||||||||||||||||
Db          1 GACAUGAACAAGAGAUGAUU 20


RESULT 5
AAT72616
ID   AAT72616 standard; DNA; 30 BP.
XX
AC   AAT72616;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-89b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-89b which
CC   contains a sequence which is complementary to at least two non-contiguous
CC   regions of a hepatitis B virus (HBV) nucleic acid. The antisense
CC   oligonucleotide may be used to detect the presence of HBV in a sample.
CC   The antisense oligonucleotide, and oligonucleotides complementary to a
CC   portion of the HBV RNA, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 9 G; 6 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 9.9;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db          1 GACATGAACAAGAGATGATT 20


RESULT 6
AAT72617
ID   AAT72617 standard; DNA; 30 BP.
XX
AC   AAT72617;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-89Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        1. .20
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   1
FT                   /*tag=  c
FT                   /mod_base= gm
FT   modified_base   2
FT                   /*tag=  d
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   3
FT                   /*tag=  e
FT                   /mod_base= cm
FT   modified_base   4
FT                   /*tag=  f
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   5
FT                   /*tag=  g
FT                   /mod_base= um
FT   modified_base   6
FT                   /*tag=  h
FT                   /mod_base= gm
FT   modified_base   7
FT                   /*tag=  i
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
```

```
FT   modified_base   8
FT                   /*tag= j
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   9
FT                   /*tag= k
FT                   /mod_base= cm
FT   modified_base   10
FT                   /*tag= l
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   11
FT                   /*tag= m
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   12
FT                   /*tag= n
FT                   /mod_base= gm
FT   modified_base   13
FT                   /*tag= o
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   14
FT                   /*tag= p
FT                   /mod_base= gm
FT   modified_base   15
FT                   /*tag= q
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   16
FT                   /*tag= r
FT                   /mod_base= um
FT   modified_base   17
FT                   /*tag= s
FT                   /mod_base= gm
FT   modified_base   18
FT                   /*tag= t
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   19
FT                   /*tag= u
FT                   /mod_base= um
FT   modified_base   20
FT                   /*tag= v
FT                   /mod_base= um
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-89Mb
CC   which contains a sequence which is complementary to at least two non-
CC   contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides complementary
CC   to a portion of the HBV RNA, may be used for inhibiting HBV replication
CC   in a cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 12 A; 3 C; 9 G; 2 T; 4 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||||||||||||||||||
Db          1 GACAUGAACAAGAGAUGAUU 20


RESULT 7
ADC64742/c
ID   ADC64742 standard; RNA; 39 BP.
XX
AC   ADC64742;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Hepatitis B virus DNA polymerase related RNA oligonucleotide.
XX
KW   screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
XX
OS   Synthetic.
OS   Hepatitis B virus.
XX
PN   KR2002007891-A.
XX
PD   29-JAN-2002.
XX
PF   19-JUL-2000; 2000KR-00041420.
XX
PR   19-JUL-2000; 2000KR-00041420.
XX
PA   (MOGA-) MOGAM BIOTECHNOLOGY INST.
PA   (VIRO-) VIROGEN CO LTD.
XX
PI   Ji HJ,  Jung SI,  Kim YC,  Min MG,  Ryu WS,  Yoon GS;
XX
DR   WPI; 2003-309015/30.
XX
PT   Screening of antiviral agents by protein-priming activity of hepatitis B
PT   virus DNA polymerase.
XX
PS   Disclosure; Page 12; 13pp; Korean.
XX
CC   The present invention describes a method of screening for an antiviral
CC   agent by the protein-priming activity of hepatitis B virus (HBV) DNA
CC   polymerase. Also described is developing an antiviral agent with a high
CC   selectivity to HBV which can be used for high-throughput screening. The
CC   present sequence represents an RNA oligonucleotide which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 39 BP; 5 A; 13 C; 3 G; 0 T; 18 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 10;  Length 39;
  Best Local Similarity   80.0%;  Pred. No. 10;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||:||::
Db         37 GACATGAACAAGAGATGATT 18


RESULT 8
AAD09094/c
ID   AAD09094 standard; DNA; 87 BP.
XX
AC   AAD09094;
XX
DT   04-SEP-2001  (first entry)
```

```
XX
DE   Hepatitis B virus FR1 strain genotype G HBeAg DNA fragment.
XX
KW   HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
KW   preS2; surface antigen; HBsAg; HBX protein; vaccine; HBeAg;
KW   liver disease; hepatitis; liver cancer; HBcAg; core antigen; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200138498-A2.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US032108.
XX
PR   24-NOV-1999;   99US-0167206P.
XX
PA   (PHAR-) PHARMASSET INC.
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Schinazi R,  De Gendt S,  Van Geyt C,  Zoulim F,  Fried M;
PI   Rossau R;
XX
DR   WPI; 2001-367676/38.
XX
PT   Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT   polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT   treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
PS   Claim 6; Page 57; 84pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) strain FR1,
CC   genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
CC   PreS2 and surface antigen HBsAg) and HBX proteins. HBV genotype G nucleic
CC   acids and polypeptides are useful for diagnosing, prognosing and treating
CC   infections caused by HBV genotype G. They can be used in a vaccine to
CC   treat or prevent HBV genotype G infection. The HBV genotype G derived
CC   nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC   sample or diagnosis of HBV genotype G infection. The presence of HBV
CC   genotype G statistically correlates with the presence of liver damage
CC   and/or liver cancer in the subject. The HBV genotype G core insert
CC   peptide encoding nucleic acid is useful for designing monitoring assays
CC   to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC   and HBeAg (genotype G e antigen) in patients infected with HBV. The
CC   antibodies or antigens of HBV genotype G are useful for identifying a
CC   stage of liver disease caused by HBV genotype G. The present sequence is
CC   a hepatitis B virus (HBV) strain FR1, genotype G DNA fragment encoding e
CC   antigen (HBeAg)
XX
SQ   Sequence 87 BP; 14 A; 24 C; 17 G; 32 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 4;  Length 87;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:|||||||||:||::
Db         43 GACATGAACAAGAGATGATT 24


RESULT 9
AAD09093/c
ID   AAD09093 standard; DNA; 129 BP.
XX
AC   AAD09093;
XX
DT   04-SEP-2001  (first entry)
XX
DE   Hepatitis B virus FR1 strain genotype G DNA fragment #1.
XX
KW   HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
KW   preS2; surface antigen; HBsAg; HBX protein; vaccine; liver disease;
KW   hepatitis; liver cancer; HBcAg; core antigen; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200138498-A2.
XX
PD   31-MAY-2001.
XX
PF   21-NOV-2000; 2000WO-US032108.
XX
PR   24-NOV-1999;   99US-0167206P.
XX
PA   (PHAR-) PHARMASSET INC.
PA   (INNO-) INNOGENETICS NV.
XX
PI   Stuyver L,  Schinazi R,  De Gendt S,  Van Geyt C,  Zoulim F,  Fried M;
PI   Rossau R;
XX
DR   WPI; 2001-367676/38.
XX
PT   Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT   polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT   treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
PS   Claim 5; Page 57; 84pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) strain FR1,
CC   genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
CC   PreS2 and surface antigen HBsAg) and HBX proteins. HBV genotype G nucleic
CC   acids and polypeptides are useful for diagnosing, prognosing and treating
CC   infections caused by HBV genotype G. They can be used in a vaccine to
CC   treat or prevent HBV genotype G infection. The HBV genotype G derived
CC   nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC   sample or diagnosis of HBV genotype G infection. The presence of HBV
CC   genotype G statistically correlates with the presence of liver damage
CC   and/or liver cancer in the subject. The HBV genotype G core insert
CC   peptide encoding nucleic acid is useful for designing monitoring assays
CC   to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC   and HBeAg (genotype G e antigen) in patients infected with HBV. The
CC   antibodies or antigens of HBV genotype G are useful for identifying a
CC   stage of liver disease caused by HBV genotype G. The present sequence is
CC   a hepatitis B virus (HBV) strain FR1, genotype G DNA fragment
XX
SQ   Sequence 129 BP; 25 A; 32 C; 26 G; 46 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 4;  Length 129;
  Best Local Similarity   80.0%;  Pred. No. 11;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:|||||||||:||::
Db         43 GACATGAACAAGAGATGATT 24


RESULT 10
AAD27422/c
ID   AAD27422 standard; DNA; 639 BP.
XX
AC   AAD27422;
XX
DT   18-APR-2002  (first entry)
XX
DE   Hepatitis B virus (HBV) core antigen (HBcAg) encoding DNA #1.
XX
KW   Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
KW   prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KW   hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW   tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KW   dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KW   food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
KW   antiprotozoal; ds.
XX
OS   Hepatitis B virus.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag=  a
FT                   /product= "HBcAg"
XX
PN   WO200198333-A2.
XX
PD   27-DEC-2001.
XX
PF   22-JUN-2001; 2001WO-GB002817.
XX
PR   22-JUN-2000; 2000GB-00015308.
PR   06-OCT-2000; 2000GB-00024544.
XX
PA   (CELL-) CELLTECH PHARM LTD.
XX
PI   Page M,  Li J,  Pumpens P;
XX
DR   WPI; 2002-098223/13.
DR   P-PSDB; AAE17018.
XX
PT   New proteins comprising a modified hepatitis B core antigen, useful as a
PT   vaccine in prophylactic or therapeutic vaccination of the human or animal
PT   body, particularly against hepatitis B virus infection.
XX
PS   Disclosure; Page 38-39; 40pp; English.
XX
CC   The invention relates to modified proteins comprising hepatitis B virus
CC   (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC   repeats has been deleted and the protein comprising the C-terminal
CC   cysteine of HBcAg. The deleted region may be replaced by an epitope from
CC   a protein other than HBcAg, in which case the HBcAg acts as a carrier to
CC   present the epitope to the immune system. This chimeric protein or its
CC   nucleic acid is useful as a vaccine or in a method of prophylactic or
CC   therapeutic vaccination of the human or animal body, particularly against
CC   HBV. The nucleic acid encoding the protein may be used in gene therapy or
CC   DNA vaccination protocols. The chimeric protein or its nucleic acid may
CC   also be used as the basis of a prophylactic vaccine against a range of
CC   diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
CC   influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
CC   immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
CC   tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
CC   diarrhoea, meningitis or gonorrhea. The present sequence is a DNA
CC   encoding Hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 6;  Length 639;
  Best Local Similarity   80.0%;  Pred. No. 13;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         43 GACATGAACAAGAGATGATT 24


RESULT 11
AAD31509/c
ID   AAD31509 standard; DNA; 639 BP.
XX
AC   AAD31509;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Hepatitis B virus core antigen (HBcAg) encoding DNA.
XX
KW   Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
KW   therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
KW   polio; herpes; rabies; AIDS; foot-and-mouth disease; ds.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag=  a
FT                   /product= "HBc protein"
FT   sig_peptide     1. .87
FT                   /*tag=  b
FT   mat_peptide     88. .636
FT                   /*tag=  c
FT                   /product= "Mature HBc protein"
XX
PN   WO200177158-A1.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-GB001607.
XX
PR   07-APR-2000; 2000EP-00107118.
XX
PA   (MEDE-) MEDEVA EURO LTD.
XX
PI   Gehin A,  Gilbert R,  Stuart D,  Rowlands D;
XX
DR   WPI; 2002-239995/29.
DR   P-PSDB; AAE19793.
XX
PT   Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for the
PT   prophylactic or therapeutic treatment of humans or animals against e.g.
PT   HB virus, viral hepatitis, hepatitis C virus, influenza, or foot-and-
PT   mouth disease.
XX
PS   Disclosure; Page 23-24; 27pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) core antigen
CC   (HBcAg) fusion proteins and polynucleotides encoding such proteins.
CC   Sequences of the invention are useful in methods of prophylactic or
CC   therapeutic vaccination or to manufacture medicaments for prophylactic or
CC   therapeutic vaccination of the human or animal body against HBV, e.g.
CC   against viral hepatitis. They are also useful as a prophylactic vaccine
CC   against e.g. hepatitis c virus, influenza, polio, herpes, rabies,
CC   acquired immune deficiency syndrome (AIDS) or foot-and-mouth disease. The
CC   present sequence is a DNA encoding hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 6;  Length 639;
  Best Local Similarity   80.0%;  Pred. No. 13;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              ||||:||||||||||:||::
Db         43 GACATGAACAAGAGATGATT 24


RESULT 12
AAH77569/c
ID   AAH77569 standard; DNA; 655 BP.
XX
AC   AAH77569;
XX
DT   19-OCT-2001  (first entry)
XX
DE   HBV genotype G strain US1 preCore/Core DNA.
XX
KW   Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW   HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW   HBeAg; ds.
XX
OS   Hepatitis B virus.
XX
PN   WO200140279-A2.
XX
PD   07-JUN-2001.
XX
PF   20-NOV-2000; 2000WO-EP011526.
```

XX
PR 03-DEC-1999; 99EP-00870252.
PR 07-DEC-1999; 99US-0169287P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.
XX
PS Claim 3; Fig 7; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC This genotype was found with a high prevalence in patients chronically
CC infected with HBV and residing in Europe and the USA. The invention
CC relates to a fully defined sequence of 3248 nucleotides as given in
CC specification, a sequence with 92% identity to the given sequence, or
CC sequence that is degenerate to the mentioned sequences. These
CC polynucleotides are useful for HBV genotyping. The proteins encoded by
CC the polynucleotides are useful for detecting antibodies in a biological
CC sample. Ligands that bind to the proteins and antibodies directed against
CC the proteins are useful for detecting the proteins and for detecting
CC HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC preparing a vaccine or medicament for treating HBV infections. The
CC present sequence is provided in an alignment of preCore/Core sequences of
CC an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC US6, US7, US9, US10) of HBV genotype G
XX
SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 20; DB 4; Length 655;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAGAGAUGAUU 20
||||:||||||||||:||::
Db 43 GACATGAACAAGAGATGATT 24

RESULT 13
AAH77568/c
ID AAH77568 standard; DNA; 655 BP.
XX
AC AAH77568;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype G strain FR2 preCore/Core DNA.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW HBeAg; ds.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP011526.
XX
PR 03-DEC-1999; 99EP-00870252.
PR 07-DEC-1999; 99US-0169287P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.
XX
PS Claim 3; Fig 7; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC This genotype was found with a high prevalence in patients chronically
CC infected with HBV and residing in Europe and the USA. The invention
CC relates to a fully defined sequence of 3248 nucleotides as given in
CC specification, a sequence with 92% identity to the given sequence, or
CC sequence that is degenerate to the mentioned sequences. These
CC polynucleotides are useful for HBV genotyping. The proteins encoded by
CC the polynucleotides are useful for detecting antibodies in a biological
CC sample. Ligands that bind to the proteins and antibodies directed against
CC the proteins are useful for detecting the proteins and for detecting
CC HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC preparing a vaccine or medicament for treating HBV infections. The
CC present sequence is provided in an alignment of preCore/Core sequences of
CC an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC US6, US7, US9, US10) of HBV genotype G
XX
SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 20; DB 4; Length 655;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAGAGAUGAUU 20
||||:||||||||||:||::
Db 43 GACATGAACAAGAGATGATT 24

RESULT 14
AAH77574/c
ID AAH77574 standard; DNA; 655 BP.
XX
AC AAH77574;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype G strain US10 preCore/Core DNA.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW HBeAg; ds.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP011526.
XX
PR 03-DEC-1999; 99EP-00870252.
PR 07-DEC-1999; 99US-0169287P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.

XX
PS Claim 3; Fig 7; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC This genotype was found with a high prevalence in patients chronically
CC infected with HBV and residing in Europe and the USA. The invention
CC relates to a fully defined sequence of 3248 nucleotides as given in
CC specification, a sequence with 92% identity to the given sequence, or
CC sequence that is degenerate to the mentioned sequences. These
CC polynucleotides are useful for HBV genotyping. The proteins encoded by
CC the polynucleotides are useful for detecting antibodies in a biological
CC sample. Ligands that bind to the proteins and antibodies directed against
CC the proteins are useful for detecting the proteins and for detecting
CC HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC preparing a vaccine or medicament for treating HBV infections. The
CC present sequence is provided in an alignment of preCore/Core sequences of
CC an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC US6, US7, US9, US10) of HBV genotype G
XX
SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 20; DB 4; Length 655;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAGAGAUGAUU 20
||||:||||||||||:||::
Db 43 GACATGAACAAGAGATGATT 24

RESULT 15
AAH77573/c
ID AAH77573 standard; DNA; 655 BP.
XX
AC AAH77573;
XX
DT 19-OCT-2001 (first entry)
XX
DE HBV genotype G strain US7 preCore/Core DNA.
XX
KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
KW HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
KW HBeAg; ds.
XX
OS Hepatitis B virus.
XX
PN WO200140279-A2.
XX
PD 07-JUN-2001.
XX
PF 20-NOV-2000; 2000WO-EP011526.
XX
PR 03-DEC-1999; 99EP-00870252.
PR 07-DEC-1999; 99US-0169287P.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Stuyver L, Van Geyt C, De Gendt S;
XX
DR WPI; 2001-374785/39.
XX
PT Novel isolated and/or purified hepatitis B virus polypeptide and
PT polynucleotide sequences that are phylogenetically different from HBV
PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
PT therapy.
XX
PS Claim 3; Fig 7; 94pp; English.
XX
CC The invention relates to the complete nucleic acid sequence of a new
CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC This genotype was found with a high prevalence in patients chronically
CC infected with HBV and residing in Europe and the USA. The invention
CC relates to a fully defined sequence of 3248 nucleotides as given in
CC specification, a sequence with 92% identity to the given sequence, or
CC sequence that is degenerate to the mentioned sequences. These
CC polynucleotides are useful for HBV genotyping. The proteins encoded by
CC the polynucleotides are useful for detecting antibodies in a biological
CC sample. Ligands that bind to the proteins and antibodies directed against
CC the proteins are useful for detecting the proteins and for detecting
CC HBcAg and HBeAg (precore precursor proteins). They are also useful for
CC preparing a vaccine or medicament for treating HBV infections. The
CC present sequence is provided in an alignment of preCore/Core sequences of
CC an HBV genotype A strain (HBVXCPS) and 7 strains (FR1, FR2, US1, US3,
CC US6, US7, US9, US10) of HBV genotype G
XX
SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 20; DB 4; Length 655;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAGAGAUGAUU 20
||||:||||||||||:||::
Db 43 GACATGAACAAGAGATGATT 24

Search completed: March 17, 2005, 06:48:42
Job time : 172.333 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-61
Perfect score: 20
Sequence: 1 gacaugaacaagagaugauu 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	18.4	92.0	763	9	CG381974	CG381974 OG1BK24TH
c	2	18.4	92.0	773	9	CG381984	CG381984 OG1BK24TV
c	3	18.4	92.0	846	9	CG373212	CG373212 OG1CZ28TH
	4	18.4	92.0	892	9	CG373225	CG373225 OG1CZ28TV
c	5	17.4	87.0	280	9	BX288914	BX288914 Arabidops
	6	17.4	87.0	305	6	CA520799	CA520799 KS11017B1
c	7	17.4	87.0	313	7	CF906991	CF906991 A0504F09-
	8	17.4	87.0	423	5	BX837806	BX837806 BX837806
c	9	17.4	87.0	426	8	AQ183243	AQ183243 HS_3140_B
	10	17.4	87.0	442	7	CN958899	CN958899 6399_1001
	11	17.4	87.0	450	8	AZ654527	AZ654527 1M0528D16
	12	17.4	87.0	509	3	AK015232	AK015232 Mus muscu
c	13	17.4	87.0	527	8	AQ961278	AQ961278 LERFK20TR
	14	17.4	87.0	570	7	CF198522	CF198522 EST0117 T
c	15	17.4	87.0	574	7	CF909462	CF909462 A0536H03-
	16	17.4	87.0	616	7	CK517114	CK517114 rswjb0_00
c	17	17.4	87.0	628	8	BH366000	BH366000 CH230-112
	18	17.4	87.0	631	2	BE388774	BE388774 601283896
c	19	17.4	87.0	652	8	BZ898091	BZ898091 CH240_13N
	20	17.4	87.0	682	9	CE163565	CE163565 tigr-gss-
c	21	17.4	87.0	683	7	CO817689	CO817689 FA_SEa001
	22	17.4	87.0	721	8	AQ961277	AQ961277 LERFK20TF
c	23	17.4	87.0	794	9	CL809904	CL809904 OR_CBa002
	24	17.4	87.0	821	2	BF678287	BF678287 602084906
c	25	17.4	87.0	859	8	CC090167	CC090167 CSU-K33r.
	26	17.4	87.0	870	8	CC131380	CC131380 NDL.76K8.
c	27	17.4	87.0	885	2	BF541940	BF541940 602068444
c	28	17.4	87.0	903	8	CC068329	CC068329 CSU-K33r.
	29	17.4	87.0	1340	3	CNS0A5V5	BX823129 Arabidops
	30	17	85.0	472	2	AW760013	AW760013 sl56h09.y
c	31	17	85.0	532	1	AL819446	AL819446 AL819446
c	32	17	85.0	577	4	BI378081	BI378081 BFLG3_001
	33	17	85.0	687	9	AG140746	AG140746 Pan trogl
	34	17	85.0	730	5	BX114353	BX114353 BX114353
c	35	17	85.0	768	9	CC786410	CC786410 ZMMBBb015
	36	17	85.0	778	7	CO368799	CO368799 RTK1_42_H
	37	17	85.0	808	9	CL543431	CL543431 OB__Ba006
	38	17	85.0	1013	9	CNS06RAQ	AL411720 T3 end of
	39	16.8	84.0	124	4	BI127956	BI127956 G068P81Y
	40	16.8	84.0	170	4	BI128183	BI128183 G072P24Y
	41	16.8	84.0	195	4	BG125265	BG125265 EST470911
	42	16.8	84.0	195	4	BG733602	BG733602 Cs-20d Ci
c	43	16.8	84.0	218	2	BE428564	BE428564 MTD008.D0
c	44	16.8	84.0	234	4	BI473621	BI473621 fp39h03.y
	45	16.8	84.0	251	8	BZ385056	BZ385056 SALK_1363

ALIGNMENTS

RESULT 1
CG381974
LOCUS CG381974 763 bp DNA linear GSS 26-AUG-2003
DEFINITION OG1BK24TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0724C23, genomic survey sequence.
ACCESSION CG381974
VERSION CG381974.1 GI:34299241
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 763)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OG1BK24TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .763
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0724C23"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 763;
Best Local Similarity 75.0%; Pred. No. 4.2e+02;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GACAUGAACAAGAGAUGAUU 20
              | ||:|||||||||:||::
Db        572 GGCATGAACAAGAGATGATT 591
```

```
RESULT 2
CG381984/c
LOCUS       CG381984                773 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OG1BK24TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0724C23,
            genomic survey sequence.
ACCESSION   CG381984
VERSION     CG381984.1  GI:34299251
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1BK24TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .773
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0724C23"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 9;  Length 773;
  Best Local Similarity   75.0%;  Pred. No. 4.3e+02;
  Matches   15;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              | ||:||||||||||:||::
Db        668 GGCATGAACAAGAGATGATT 649


RESULT 3
CG373212/c
LOCUS       CG373212                846 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OG1CZ28TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0734E08,
            genomic survey sequence.
ACCESSION   CG373212
VERSION     CG373212.1  GI:34290479
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1CZ28TV
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0734E08"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 9;  Length 846;
  Best Local Similarity   75.0%;  Pred. No. 4.3e+02;
  Matches   15;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              | ||:||||||||||:||::
Db        540 GGCATGAACAAGAGATGATT 521


RESULT 4
CG373225
LOCUS       CG373225                892 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OG1CZ28TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0734E08,
            genomic survey sequence.
ACCESSION   CG373225
VERSION     CG373225.1  GI:34290492
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1CZ28TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .892
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0734E08"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 9;  Length 892;
  Best Local Similarity   75.0%;  Pred. No. 4.4e+02;
  Matches   15;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAUU 20
              | ||:||||||||||:||::
```

```
Db        353 GGCATGAACAAGAGATGATT 372


RESULT 5
BX288914/c
LOCUS       BX288914                 280 bp    DNA     linear   GSS 02-APR-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-421D09-018141,
            genomic survey sequence.
ACCESSION   BX288914
VERSION     BX288914.1  GI:28887910
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
  TITLE     GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
  JOURNAL   Bioinformatics 19 (11), 1441-1442 (2003)
  MEDLINE   22755829
   PUBMED   12874060
REFERENCE   2
  AUTHORS   Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
            Weisshaar,B.
  TITLE     An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
            flanking sequence tag-based reverse genetics
  JOURNAL   Plant Mol. Biol. 53 (1-2), 247-259 (2003)
  MEDLINE   23117147
   PUBMED   14756321
REFERENCE   3
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
            Weisshaar,B.
  TITLE     High-throughput generation of sequence indexes from T-DNA
            mutagenized Arabidopsis thaliana lines
  JOURNAL   BioTechniques 35 (6), 1164-1168 (2003)
   PUBMED   14682050
REFERENCE   4  (bases 1 to 280)
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence has been recovered from the left border of the T-DNA.
            It indicates an insertion within the locus defined by BAC clone
            f3k23. Details on the protocols used for generation of the sequence
            are described in References 1-3. The sequences are generated at the
            MPI for Plant Breeding Research in the context of the GABI-Kat
            project. GABI-Kat is part of the German Plant Genomics program
            designated 'GABI'. Information on line availability can be found
            at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .280
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-421D09-018141"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 280;
  Best Local Similarity   78.9%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 ACAUGAACAAGAGAUGAUU 20
              |||:||||||||| ||::
Db         80 ACATGAACAAGAGAAGATT 62


RESULT 6
CA520799
LOCUS       CA520799                 305 bp    mRNA    linear   EST 15-NOV-2002
DEFINITION  KS11017B12 KS11 Capsicum annuum cDNA, mRNA sequence.
ACCESSION   CA520799
VERSION     CA520799.1  GI:25034824
KEYWORDS    EST.
SOURCE      Capsicum annuum
  ORGANISM  Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Capsicum.
REFERENCE   1  (bases 1 to 305)
  AUTHORS   Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S.,
            Hur,C.-G. and Choi,D.
  TITLE     Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
            annuum L.) and Sequence Analysis in Relation to Hypersensitive
            Response Against Pathogen
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Doil Choi
            Genome Research Center and National Center for Genome Information
            Korea Research Institute of Bioscience and Biotechnology
            P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea
            Tel: 82-42-860-4340
            Fax: 82-42-860-4309
            Email: doil@mail.kribb.re.kr
            Plate: 017  row: B  column: 12.
FEATURES             Location/Qualifiers
     source          1. .305
                     /organism="Capsicum annuum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4072"
                     /clone_lib="KS11"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 6;  Length 305;
  Best Local Similarity   78.9%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAU 19
              ||||:|||||||||: |:
Db        212 GACATGAACAAGAGATCAT 230


RESULT 7
CF906991/c
LOCUS       CF906991                 313 bp    mRNA    linear   EST 05-NOV-2003
DEFINITION  A0504F09-5 NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)
            Mus musculus cDNA clone NIA:A0504F09 IMAGE:30743204 5', mRNA
            sequence.
ACCESSION   CF906991
VERSION     CF906991.1  GI:38177928
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
  TITLE     Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
  JOURNAL   Genome Res. 11 (9), 1553-1558 (2001)
  MEDLINE   21429098
   PUBMED   11544199
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
```

National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: A0504 row: F column: 09
Seq primer: M13 Reverse
High quality sequence stop: 313
POLYA=No.

FEATURES Location/Qualifiers
source 1. .313
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H/He mice"
/db_xref="niaEST:A0504F09-5"
/db_xref="taxon:10090"
/clone="NIA:A0504F09 IMAGE:30743204"
/dev_stage="9-15C cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Akihiro Umezawa (Keio University School of Medicine, Japan). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from 2.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 313;
Best Local Similarity 78.9%; Pred. No. 1.1e+03;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GACAUGAACAAGAGAUGAU 19
              || |:||||||||||:||:
Db        289 GAAATGAACAAGAGATGAT 271
```

RESULT 8
BX837806
LOCUS BX837806 423 bp mRNA linear EST 11-FEB-2004
DEFINITION BX837806 Arabidopsis thaliana Hormone Treated Callus Col-0 Arabidopsis thaliana cDNA clone GSLTPGH21ZC04 5PRIM, mRNA sequence.
ACCESSION BX837806
VERSION BX837806.1 GI:42531889
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 423)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL Unpublished (2004)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length librairies construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES Location/Qualifiers
source 1. .423
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH21ZC04"
/tissue_type="Hormone Treated Callus"
/clone_lib="Arabidopsis thaliana Hormone Treated Callus Col-0"

ORIGIN

Query Match 87.0%; Score 17.4; DB 5; Length 423;
Best Local Similarity 73.7%; Pred. No. 1.2e+03;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 ACAUGAACAAGAGAUGAUU 20
              |||:||| ||||||:||::
Db        403 ACATGAAGAAGAGATGATT 421
```

RESULT 9
AQ183243/c
LOCUS AQ183243 426 bp DNA linear GSS 01-NOV-1998
DEFINITION HS_3140_B2_B02_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3140 Col=4 Row=D, genomic survey sequence.
ACCESSION AQ183243
VERSION AQ183243.1 GI:3580610
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 426)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3140 row: D column: 4
Class: BAC ends
High quality sequence stop: 426.

FEATURES Location/Qualifiers

```
     source          1. .426
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3140 Col=4 Row=D"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 8;  Length 426;
  Best Local Similarity   78.9%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 ACAUGAACAAGAGAUGAUU 20
              |||:||||||||||:|| :
Db        273 ACATGAACAAGAGATGACT 255


RESULT 10
CN958899
LOCUS       CN958899                442 bp    mRNA    linear   EST 08-JUN-2004
DEFINITION  6399_100122_44 Fundulus Heteroclitus Liver Fundulus heteroclitus
            cDNA, mRNA sequence.
ACCESSION   CN958899
VERSION     CN958899.1  GI:48440488
KEYWORDS    EST.
SOURCE      Fundulus heteroclitus (killifish)
  ORGANISM  Fundulus heteroclitus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Cyprinodontiformes; Fundulidae; Fundulus.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   Crawford,D.L., Oleksiak,M.F., Kolell,K.J., Paschall,J., VanWye,J.,
            Roach,J.L. and Whitehead,J.A.
  TITLE     Fundulus Functional Genomics: EST Database for Teleost Fish
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Crawford, Douglas L.
            Marine Genomics - Crawford Lab
            Rosenstiel School of Marine and Atmospheric Science - University of
            Miami
            4600 Rickenbacker Causeway, Miami, FL 33149-1098 USA
            Tel: 305 361 4121
            Email: dcrawford@rsmas.Miami.edu
            Database Web Interface
            http://genomics.rsmas.miami.edu/funnybase/super_craw3/
            Plate: 100122  row: E  column: 6.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Fundulus heteroclitus"
                     /mol_type="mRNA"
                     /db_xref="taxon:8078"
                     /tissue_type="Liver"
                     /clone_lib="Fundulus Heteroclitus Liver"
                     /note="Organ: Liver"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 7;  Length 442;
  Best Local Similarity   78.9%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAU 19
              | ||:||||||||||:||:
Db        201 GTCATGAACAAGAGATGAT 219


RESULT 11
AZ654527
LOCUS       AZ654527                450 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0528D16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0528D16 R, genomic survey sequence.
ACCESSION   AZ654527
VERSION     AZ654527.1  GI:11791673
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 450)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0528  row: D  column: 16
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .450
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0528D16"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 8;  Length 450;
  Best Local Similarity   78.9%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GACAUGAACAAGAGAUGAU 19
              || |:||||||||||:||:
Db        132 GAAATGAACAAGAGATGAT 150


RESULT 12
AK015232
LOCUS       AK015232                509 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4930429C20 product:unclassifiable, full insert
```

sequence.
ACCESSION AK015232
VERSION AK015232.1 GI:12853489
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 509)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
FEATURES Location/Qualifiers
source 1. .509
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4930429C20"
/db_xref="taxon:10090"
/clone="4930429C20"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature 1. .509
/note="unclassifiable"
ORIGIN

Query Match 87.0%; Score 17.4; DB 3; Length 509;
Best Local Similarity 73.7%; Pred. No. 1.2e+03;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ACAUGAACAAGAGAUGAUU 20
| |:|||||||||:||::
Db 443 AGATGAACAAGAGATGATT 461

RESULT 13
AQ961278/c
LOCUS AQ961278 527 bp DNA linear GSS 28-JAN-2000
DEFINITION LERFK20TR LERA Arabidopsis thaliana genomic clone LERFK20, genomic survey sequence.
ACCESSION AQ961278
VERSION AQ961278.1 GI:6788979
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 527)
AUTHORS Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T., Feldblyum,T., Liang,F., Creasy,T. and Fraser,C.M.
TITLE Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
JOURNAL Unpublished (2000)
COMMENT Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: at@tigr.org
For addtional information, see http://www.tigr.org/tdb/at/at.html
Seq primer: TR
Class: shotgun.
FEATURES Location/Qualifiers
source 1. .527
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LERFK20"
/clone_lib="LERA"
/note="Organ: Leaf; Vector: pHOS1; Total genomic DNA was

sheared to 0.9-1 Kbp before ligation."
ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 527;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 ACAUGAACAAGAGAUGAUU 20
              |||:|||||||||| ||::
Db        160 ACATGAACAAGAGAAGATT 142
```

RESULT 14
CF198522
LOCUS CF198522 570 bp mRNA linear EST 01-AUG-2003
DEFINITION EST0117 Tamarix androssowii leaf Tamarix androssowii cDNA, mRNA sequence.
ACCESSION CF198522
VERSION CF198522.1 GI:33392895
KEYWORDS EST.
SOURCE Tamarix androssowii
ORGANISM Tamarix androssowii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Tamaricaceae; Tamarix.
REFERENCE 1 (bases 1 to 570)
AUTHORS Wang,Y., Yang,C., Jiang,J., Liu,G., Wu,J. and Liu,Z.
TITLE EST aquired from cDNA library of Tamarix androssowii treated with NaHCO3
JOURNAL Unpublished (2003)
COMMENT Contact: Yucheng Wang
Forestry Source and Environment College
Northeast Forestry University
Hexing 26, Harbin, Heilongjiang, 150040, P.R. China
Tel: 086-451-2190607
Email: WANGYUCHENG1029@YAHOO.COM.CN.
FEATURES Location/Qualifiers
source 1. .570
/organism="Tamarix androssowii"
/mol_type="mRNA"
/db_xref="taxon:189785"
/tissue_type="leaf"
/clone_lib="Tamarix androssowii leaf"
ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 570;
Best Local Similarity 73.7%; Pred. No. 1.2e+03;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 ACAUGAACAAGAGAUGAUU 20
              |||:|||||||||| :||::
Db        143 ACATGAACAAGAGGTGATT 161
```

RESULT 15
CF909462/c
LOCUS CF909462 574 bp mRNA linear EST 05-NOV-2003
DEFINITION A0536H03-5 NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1) Mus musculus cDNA clone NIA:A0536H03 IMAGE:30746294 5', mRNA sequence.
ACCESSION CF909462
VERSION CF909462.1 GI:38180399
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 574)
AUTHORS Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
TITLE Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method
JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098
PUBMED 11544199
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: A0536 row: H column: 03
Seq primer: M13 Reverse
High quality sequence stop: 574
POLYA=No.
FEATURES Location/Qualifiers
source 1. .574
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H/He mice"
/db_xref="niaEST:A0536H03-5"
/db_xref="taxon:10090"
/clone="NIA:A0536H03 IMAGE:30746294"
/dev_stage="9-15C cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Akihiro Umezawa (Keio University School of Medicine, Japan). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from 2.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."
ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 574;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GACAUGAACAAGAGAUGAU 19
              || |:|||||||||:||:
Db        510 GAAATGAACAAGAGATGAT 492
```

Search completed: March 17, 2005, 11:07:46
Job time : 1386.27 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-62
Perfect score: 20
Sequence: 1 taagggtcgauguccaugcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	6	AR027817	AR027817 Sequence
2		20	100.0	30	6	AR027842	AR027842 Sequence
3		20	100.0	30	6	AR027843	AR027843 Sequence
c 4		20	100.0	81	6	I92348	I92348 Sequence 9
c 5		20	100.0	174	14	S77749	S77749 preC (X/pre
c 6		20	100.0	253	14	AY329529	AY329529 Hepatitis
c 7		20	100.0	253	14	AY329562	AY329562 Hepatitis
c 8		20	100.0	253	14	AY329568	AY329568 Hepatitis
c 9		20	100.0	253	14	AY329573	AY329573 Hepatitis
c 10		20	100.0	253	14	AY329581	AY329581 Hepatitis
c 11		20	100.0	294	14	AF390000	AF390000 Hepatitis
c 12		20	100.0	333	14	HPBHBED	L12359 Hepatitis B
c 13		20	100.0	398	14	AB167603	AB167603 Hepatitis
c 14		20	100.0	398	14	AB167637	AB167637 Hepatitis
c 15		20	100.0	406	14	AB163815	AB163815 Hepatitis
c 16		20	100.0	406	14	AB163817	AB163817 Hepatitis
c 17		20	100.0	439	14	AY254503	AY254503 Hepatitis
c 18		20	100.0	488	14	AY274419	AY274419 Hepatitis
c 19		20	100.0	488	14	AY274420	AY274420 Hepatitis
c 20		20	100.0	488	14	AY274422	AY274422 Hepatitis
c 21		20	100.0	488	14	AY274427	AY274427 Hepatitis
c 22		20	100.0	488	14	AY274428	AY274428 Hepatitis
c 23		20	100.0	488	14	AY274429	AY274429 Hepatitis
c 24		20	100.0	488	14	AY274430	AY274430 Hepatitis
c 25		20	100.0	488	14	AY274431	AY274431 Hepatitis
c 26		20	100.0	488	14	AY274432	AY274432 Hepatitis
c 27		20	100.0	488	14	AY274433	AY274433 Hepatitis
c 28		20	100.0	488	14	AY274434	AY274434 Hepatitis
c 29		20	100.0	488	14	AY274436	AY274436 Hepatitis
c 30		20	100.0	493	14	S79556	S79556 X, preC [he
c 31		20	100.0	548	14	AY382500	AY382500 Hepatitis
c 32		20	100.0	548	14	AY382501	AY382501 Hepatitis
c 33		20	100.0	548	14	AY382502	AY382502 Hepatitis
c 34		20	100.0	548	14	AY382521	AY382521 Hepatitis
c 35		20	100.0	548	14	AY382522	AY382522 Hepatitis
c 36		20	100.0	548	14	AY382523	AY382523 Hepatitis
c 37		20	100.0	548	14	AY382524	AY382524 Hepatitis
c 38		20	100.0	548	14	AY382525	AY382525 Hepatitis
c 39		20	100.0	548	14	AY382526	AY382526 Hepatitis
c 40		20	100.0	548	14	AY382527	AY382527 Hepatitis
c 41		20	100.0	609	14	AF289954	AF289954 Hepatitis
c 42		20	100.0	609	14	AF289965	AF289965 Hepatitis
c 43		20	100.0	626	14	AY254500	AY254500 Hepatitis
c 44		20	100.0	639	6	AX278066	AX278066 Sequence
c 45		20	100.0	639	6	AX342485	AX342485 Sequence

ALIGNMENTS

RESULT 1
AR027817
LOCUS AR027817 20 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 15 from patent US 5856459.
ACCESSION AR027817
VERSION AR027817.1 GI:5938637
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and Mills,J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 15 05-JAN-1999;
FEATURES Location/Qualifiers
source 1. .20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 85.0%; Pred. No. 14;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAAGGGTCGAUGUCCAUGCC 20
||||||||||:|:|||:|||
Db 1 TAAGGGTCGATGTCCATGCC 20

RESULT 2
AR027842
LOCUS AR027842 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 40 from patent US 5856459.
ACCESSION AR027842
VERSION AR027842.1 GI:5938662
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and

```
          Mills,J.S.
  TITLE   Oligonucleotides specific for hepatitis B virus
  JOURNAL Patent: US 5856459-A 40 05-JAN-1999;
FEATURES            Location/Qualifiers
     source         1. .30
                    /organism="unknown"
                    /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 30;
  Best Local Similarity  85.0%;  Pred. No. 14;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db         11 TAAGGGTCGATGTCCATGCC 30


RESULT 3
AR027843
LOCUS       AR027843                30 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 41 from patent US 5856459.
ACCESSION   AR027843
VERSION     AR027843.1  GI:5938663
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and
            Mills,J.S.
  TITLE     Oligonucleotides specific for hepatitis B virus
  JOURNAL   Patent: US 5856459-A 41 05-JAN-1999;
FEATURES            Location/Qualifiers
     source         1. .30
                    /organism="unknown"
                    /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 30;
  Best Local Similarity  85.0%;  Pred. No. 14;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db         11 TAAGGGTCGATGTCCATGCC 30


RESULT 4
I92348/c
LOCUS       I92348                  81 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 9 from patent US 5728518.
ACCESSION   I92348
VERSION     I92348.1  GI:3936818
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Carmichael,E.
  TITLE     Antiviral poly-and oligonucleotides
  JOURNAL   Patent: US 5728518-A 9 17-MAR-1998;
FEATURES            Location/Qualifiers
     source         1. .81
                    /organism="unknown"
                    /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 81;
  Best Local Similarity  85.0%;  Pred. No. 15;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db         71 TAAGGGTCGATGTCCATGCC 52


RESULT 5
S77749/c
LOCUS       S77749                  174 bp    DNA     linear   VRL 06-MAY-2003
DEFINITION  preC {X/preC region, deletion mutant} [hepatitis B virus HBV,
            host=human, serum, patient 5 isolate, Genomic DNA Mutant, 174 nt].
ACCESSION   S77749
VERSION     S77749.1  GI:999129
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 174)
  AUTHORS   Feitelson,M.A., Duan,L.X., Guo,J. and Blumberg,B.S.
  TITLE     X region deletion mutants associated with surface antigen-positive
            hepatitis B virus infections
  JOURNAL   Gastroenterology 108 (6), 1810-1819 (1995)
  MEDLINE   95285997
   PUBMED   7768387
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 165980] from the original journal article.
FEATURES            Location/Qualifiers
     source         1. .174
                    /organism="Hepatitis B virus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10407"
     gene           69. .>174
                    /gene="preC"
                    /note="no start codon found"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 174;
  Best Local Similarity  85.0%;  Pred. No. 16;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        172 TAAGGGTCGATGTCCATGCC 153


RESULT 6
AY329529/c
LOCUS       AY329529                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate A611252E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329529
VERSION     AY329529.1  GI:37625315
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES            Location/Qualifiers
     source         1. .253
                    /organism="Hepatitis B virus"
                    /mol_type="genomic DNA"
                    /isolate="A611252E"
```

```
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95861.1"
                     /db_xref="GI:37625316"
                     /translation="STTDLEAYFKDCVFKDWEELGEETRLMIFVLGGCRHKLVCAPAS
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95862.1"
                     /db_xref="GI:37625317"
                     /translation="MQLFHLCLVISCTCPTFQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 7
AY329562/c
LOCUS       AY329562                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D273984E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329562
VERSION     AY329562.1  GI:37625413
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D273984E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95926.1"
                     /db_xref="GI:37625414"
                     /translation="STTDLEAYFKDCLFKDWEELGEEIRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95927.1"
                     /db_xref="GI:37625415"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 8
AY329568/c
LOCUS       AY329568                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D296668E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329568
VERSION     AY329568.1  GI:37625431
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D296668E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95938.1"
                     /db_xref="GI:37625432"
                     /translation="STTDLEAYFKDCLFKDWEELGEELRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95939.1"
                     /db_xref="GI:37625433"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 9
AY329573/c
LOCUS       AY329573                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D604917E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329573
VERSION     AY329573.1  GI:37625446
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
```

```
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D604917E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95948.1"
                     /db_xref="GI:37625447"
                     /translation="STTDLEAYFKDCLFKDWEELGEEIRLKVFVLGGCRHKLVCAPAS
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95949.1"
                     /db_xref="GI:37625448"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              |||||||||||:|:|||:|||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 10
AY329581/c
LOCUS       AY329581                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D639472E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329581
VERSION     AY329581.1  GI:37625470
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D639472E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95964.1"
                     /db_xref="GI:37625471"
                     /translation="STTDLEAYFKDCLFKDWEELGEELRLLIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95965.1"
                     /db_xref="GI:37625472"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              |||||||||||:|:|||:|||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 11
AF390000/c
LOCUS       AF390000                 294 bp    DNA     linear   VRL 06-MAR-2002
DEFINITION  Hepatitis B virus isolate D3 X protein gene, partial cds; and
            nonfunctional precore/core protein gene, partial sequence.
ACCESSION   AF390000
VERSION     AF390000.1  GI:16266099
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   Castro,L.D., Niel,C. and Gomes,S.A.
  TITLE     Low frequency of mutations in the core promoter and precore regions
            of hepatitis B virus in anti-HBe positive Brazilian carriers
  JOURNAL   BMC Microbiol. 1 (1), 10 (2001)
   PUBMED   11472634
REFERENCE   2  (bases 1 to 294)
  AUTHORS   De Castro,L., Niel,C. and Gomes,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-2001) Virology, FIOCRUZ, Av. Brasil 4365, Rio de
            Janeiro, RJ 21045-900, Brazil
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D3"
                     /db_xref="taxon:10407"
     CDS             <1. .119
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAL16752.1"
                     /db_xref="GI:16266100"
                     /translation="FKDWEELGDDSRLMIYVLGGCRHKLVCAPAPCNFFTSA"
     misc_feature    95. .>294
                     /note="nonfunctional precore/core protein due to mutation"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 294;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              |||||||||||:|:|||:|||
Db        198 TAAGGGTCGATGTCCATGCC 179
```

```
RESULT 12
HPBHBED/c
LOCUS       HPBHBED                  333 bp    DNA     linear   VRL 28-JUN-1996
DEFINITION  Hepatitis B virus precore and core protein gene, 5' end of cds.
ACCESSION   L12359
VERSION     L12359.1  GI:306267
KEYWORDS    HBcAg protein; HBeAg protein; core protein; nucleotide binding
            protein; precore protein.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (sites)
  AUTHORS   Tong,S.P., Li,J.S., Vitvitski,L. and Trepo,C.
  TITLE     Active hepatitis B virus replication in the presence of anti-HBe is
            associated with viral variants containing an inactive pre-C region
  JOURNAL   Virology 176 (2), 596-603 (1990)
  MEDLINE   90266476
   PUBMED   2345966
REFERENCE   2  (bases 1 to 333)
  AUTHORS   Li,J.S., Tong,S.P., Wen,Y.M., Vitvitski,L., Zhang,Q. and Trepo,C.
  TITLE     Hepatitis B virus genotype A rarely circulates as an HBe-minus
            mutant: possible contribution of a single nucleotide in the precore
            region
  JOURNAL   J. Virol. 67 (9), 5402-5410 (1993)
  MEDLINE   93353617
   PUBMED   8350403
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
                     /note="genotype D; from French HBeAg- patient"
     CDS             1. .>333
                     /standard_name="HBeAg"
                     /pseudo
                     /codon_start=1
                     /product="precore protein"
     variation       82. .84
                     /note="HBe-abolishing"
     CDS             88. .>333
                     /standard_name="HBcAb"
                     /codon_start=1
                     /protein_id="AAB03102.1"
                     /db_xref="GI:306268"
                     /translation="MDIDPYKEFGATVELLTFLPSDFFPSVRDLLDTAAALYREALES
                     PEHCSHHHTALRQAILCWGELMTLATWVGANLDDPASR"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 333;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        104 TAAGGGTCGATGTCCATGCC 85


RESULT 13
AB167603/c
LOCUS       AB167603                 398 bp    DNA     linear   VRL 01-OCT-2004
DEFINITION  Hepatitis B virus gene for polyprotein, partial cds, clone: NEP75.
ACCESSION   AB167603
VERSION     AB167603.1  GI:53148166
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1
  AUTHORS   Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
  TITLE     A Case-control Study for Differences among Hepatitis B Virus
            Infections of Genotypes A (Subtypes Aa and Ae) and D
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 398)
  AUTHORS   Tanaka,Y. and Mizokami,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-2004) Yasuhito Tanaka, Nagoya City University
            Graduate School of Medical Sciences, Department of Clinical
            Molecular Informative Medicine; 1 Kawasumi, Mizuho-cho, Mizuho-ku,
            Nagoya, Aichi 467-8601, Japan (E-mail:ytanaka@med.nagoya-cu.ac.jp,
            Tel:81-52-853-8292, Fax:81-52-842-0021)
FEATURES             Location/Qualifiers
     source          1. .398
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
                     /clone="NEP75"
     CDS             184. .>398
                     /note="precore-core region"
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="BAD52175.1"
                     /db_xref="GI:53148167"
                     /translation="MQLFHLCLIISCTCPTFQASKLCLGWLWGMDIDPYKEFGATVEL
                     LSFLPSDFFPSVRDLLDTASALYREAL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 398;
  Best Local Similarity   85.0%;  Pred. No. 18;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        287 TAAGGGTCGATGTCCATGCC 268


RESULT 14
AB167637/c
LOCUS       AB167637                 398 bp    DNA     linear   VRL 01-OCT-2004
DEFINITION  Hepatitis B virus gene for polyprotein, partial cds, clone: NEP27.
ACCESSION   AB167637
VERSION     AB167637.1  GI:53148360
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1
  AUTHORS   Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
  TITLE     A Case-control Study for Differences among Hepatitis B Virus
            Infections of Genotypes A (Subtypes Aa and Ae) and D
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 398)
  AUTHORS   Tanaka,Y. and Mizokami,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-2004) Yasuhito Tanaka, Nagoya City University
            Graduate School of Medical Sciences, Department of Clinical
            Molecular Informative Medicine; 1 Kawasumi, Mizuho-cho, Mizuho-ku,
            Nagoya, Aichi 467-8601, Japan (E-mail:ytanaka@med.nagoya-cu.ac.jp,
            Tel:81-52-853-8292, Fax:81-52-842-0021)
FEATURES             Location/Qualifiers
     source          1. .398
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
                     /clone="NEP27"
     CDS             184. .>398
                     /note="precore-core region"
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="BAD52209.1"
                     /db_xref="GI:53148361"
                     /translation="MQLFHLCLIISCTCPTFQASKLCLGWLWGMDIDPYKEFGATVEL
                     LSFLPSDFFPSVRDLLDTASALYREAL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 398;
```

```
  Best Local Similarity   85.0%;  Pred. No. 18;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        287 TAAGGGTCGATGTCCATGCC 268


RESULT 15
AB163815/c
LOCUS       AB163815                406 bp    DNA     linear   VRL 29-JUN-2004
DEFINITION  Hepatitis B virus X, preC/C genes for polyproteins, isolate:
            SAF662.
ACCESSION   AB163815
VERSION     AB163815.1  GI:49387444
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1
  AUTHORS   Hasegawa,I., Tanaka,Y., Kramvis,A., Kato,T., Sugauchi,F.,
            Acharya,S.K., Orito,E., Ueda,R., Kew,M.C. and Mizokami,M.
  TITLE     Novel hepatitis B virus genotype a subtyping assay that
            distinguishes subtype aa from ae and its application in
            epidemiological studies
  JOURNAL   J. Virol. 78 (14), 7575-7581 (2004)
   PUBMED   15220432
REFERENCE   2  (bases 1 to 406)
  AUTHORS   Hasegawa,I., Tanaka,Y. and Mizokami,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-2004) Izumi Hasegawa, Nagoya City University
            Graduate School, Department of Internal Medicine and Molecular
            Science; 1 Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya, Aichi 467-8601,
            Japan (E-mail:izu-hase@med.nagoya-cu.ac.jp, Tel:81-52-853-8216,
            Fax:81-52-852-0849)
FEATURES             Location/Qualifiers
     source          1. .406
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="SAF662"
                     /db_xref="taxon:10407"
     gene            1. .208
                     /gene="X"
     CDS             <1. .208
                     /gene="X"
                     /codon_start=2
                     /product="polyprotein"
                     /protein_id="BAD24924.1"
                     /db_xref="GI:49387445"
                     /translation="QILPKVLHKRTLGLPAMSTTDLEAYFKDCVFKDWEELGEETRLM
                     IFVLGGCRHKLVCAPSSCNFFTSA"
     gene            184. .406
                     /gene="preC/C"
     CDS             184. .>406
                     /gene="preC/C"
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="BAD24925.1"
                     /db_xref="GI:49387446"
                     /translation="MQLFHLCLIISCTCPTVQASKLCLGWLWGMDIDPYKEFGATVEL
                     LSFLPQDFFPSVRDLLDTASALYREALESP"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 406;
  Best Local Similarity   85.0%;  Pred. No. 18;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        287 TAAGGGTCGATGTCCATGCC 268


Search completed: March 17, 2005, 08:14:17

Job time : 684.733 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-62
Perfect score: 20
Sequence: 1 taagggtcgauguccaugcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	2	AAT72572	Aat72572 Hepatitis
	2	20	100.0	20	2	AAT72573	Aat72573 Hepatitis
	3	20	100.0	30	2	AAT72619	Aat72619 Hepatitis
	4	20	100.0	30	2	AAT72620	Aat72620 Hepatitis
	5	20	100.0	30	2	AAT72618	Aat72618 Hepatitis
	6	20	100.0	30	2	AAT72621	Aat72621 Hepatitis
c	7	20	100.0	31	10	ADC64743	Adc64743 Hepatitis
c	8	20	100.0	639	6	AAD27422	Aad27422 Hepatitis
c	9	20	100.0	639	6	AAD31509	Aad31509 Hepatitis
c	10	20	100.0	663	3	AAA71734	Aaa71734 HBV fusio
c	11	20	100.0	669	12	ADO07220	Ado07220 Hepatitis
c	12	20	100.0	1334	2	AAV82691	Aav82691 Fulminant
c	13	20	100.0	1395	2	AAV82688	Aav82688 Fulminant
c	14	20	100.0	1400	2	AAV82687	Aav82687 Fulminant
c	15	20	100.0	1445	2	AAV82692	Aav82692 Fulminant
c	16	20	100.0	1445	2	AAV82685	Aav82685 Fulminant
c	17	20	100.0	1445	2	AAV82690	Aav82690 Fulminant
c	18	20	100.0	1445	2	AAV82684	Aav82684 Fulminant
c	19	20	100.0	1500	2	AAV82686	Aav82686 Fulminant
c	20	20	100.0	1500	2	AAV82689	Aav82689 Fulminant
c	21	20	100.0	2342	1	AAN93072	Aan93072 Sequence
c	22	20	100.0	3182	6	AAD31765	Aad31765 Hepatitis
c	23	20	100.0	3182	9	ACA62422	Aca62422 Hepatitis
c	24	20	100.0	3182	10	AAD60866	Aad60866 Hepatitis
c	25	20	100.0	5618	2	AAQ88310	Aaq88310 Plasmid p
c	26	20	100.0	7991	6	AAS16094	Aas16094 HBV viral
c	27	20	100.0	8007	6	AAS16092	Aas16092 HBV viral
c	28	20	100.0	8717	6	AAS16093	Aas16093 HBV viral
c	29	19	95.0	34	10	ADJ94544	Adj94544 SDMCore d
	30	19	95.0	34	10	ADJ94545	Adj94545 SDMCore d
c	31	19	95.0	39	13	ADR89273	Adr89273 Lab-on-ch
c	32	19	95.0	39	13	ADR89266	Adr89266 Lab-on-ch
c	33	18.8	94.0	30	2	AAQ45813	Aaq45813 HBV ampli
c	34	18.8	94.0	30	2	AAV07810	Aav07810 HBV.D46 a
c	35	18.8	94.0	30	2	AAV83039	Aav83039 Amplifier
	36	18.8	94.0	50	2	AAQ06723	Aaq06723 :HBV.LLA2
c	37	18.4	92.0	22	10	ADG46961	Adg46961 PCR prime
c	38	18.4	92.0	22	11	ADM83206	Adm83206 PCR prime
c	39	18.4	92.0	24	6	ABK44212	Abk44212 B cell ep
c	40	18.4	92.0	24	6	ABK67439	Abk67439 Primer #1
c	41	18.4	92.0	24	6	ABK67506	Abk67506 Hepatitis
c	42	18.4	92.0	24	10	ADE80023	Ade80023 Primer fo
c	43	18.4	92.0	24	10	ADE10976	Ade10976 Chimeric
c	44	18.4	92.0	24	10	ADG46965	Adg46965 PCR prime
c	45	18.4	92.0	24	11	ADM83210	Adm83210 PCR prime

ALIGNMENTS

RESULT 1
AAT72572
ID AAT72572 standard; DNA; 20 BP.
XX
AC AAT72572;
XX
DT 03-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV92b.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1. .20
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
XX
PN WO9639502-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP002432.
XX
PR 06-JUN-1995; 95US-00467397.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (HYBR-) HYBRIDON INC.
XX
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX
DR WPI; 1997-043124/04.
XX
PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
PS Claim 1; Page 12; 81pp; English.
XX
CC The present sequence represents a synthetic oligonucleotide HBV92b which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a

```
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 4 A; 5 C; 6 G; 5 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   85.0%;  Pred. No. 2.5;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db          1 TAAGGGTCGATGTCCATGCC 20


RESULT 2
AAT72573
ID   AAT72573 standard; DNA; 20 BP.
XX
AC   AAT72573;
XX
DT   03-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV92Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .20
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        11. .20
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   11
FT                   /*tag=  c
FT                   /mod_base= um
FT   modified_base   12
FT                   /*tag=  d
FT                   /mod_base= gm
FT   modified_base   13
FT                   /*tag=  e
FT                   /mod_base= um
FT   modified_base   14
FT                   /*tag=  f
FT                   /mod_base= cm
FT   modified_base   15
FT                   /*tag=  g
FT                   /mod_base= cm
FT   modified_base   16
FT                   /*tag=  h
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   17
FT                   /*tag=  i
FT                   /mod_base= um
FT   modified_base   18
FT                   /*tag=  j
FT                   /mod_base= gm
FT   modified_base   19
FT                   /*tag=  k
FT                   /mod_base= cm
FT   modified_base   20
FT                   /*tag=  l
FT                   /mod_base= cm
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV92Mb which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 4 A; 5 C; 6 G; 2 T; 3 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||||||||||||
Db          1 TAAGGGTCGAUGUCCAUGCC 20


RESULT 3
AAT72619
ID   AAT72619 standard; DNA; 30 BP.
XX
AC   AAT72619;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-90Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        1. .10
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   1
FT                   /*tag=  c
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   2
FT                   /*tag=  d
FT                   /mod_base= gm
FT   modified_base   3
FT                   /*tag=  e
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   4
FT                   /*tag=  f
FT                   /mod_base= gm
FT   modified_base   5
FT                   /*tag=  g
```

```
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   6
FT                   /*tag=  h
FT                   /mod_base= um
FT   modified_base   7
FT                   /*tag=  i
FT                   /mod_base= gm
FT   modified_base   8
FT                   /*tag=  j
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   9
FT                   /*tag=  k
FT                   /mod_base= um
FT   modified_base   10
FT                   /*tag=  l
FT                   /mod_base= um
FT   misc_RNA        21. .30
FT                   /*tag=  m
FT                   /note= "2'-OMe RNA"
FT   modified_base   21
FT                   /*tag=  n
FT                   /mod_base= um
FT   modified_base   22
FT                   /*tag=  o
FT                   /mod_base= gm
FT   modified_base   23
FT                   /*tag=  p
FT                   /mod_base= um
FT   modified_base   24
FT                   /*tag=  q
FT                   /mod_base= cm
FT   modified_base   25
FT                   /*tag=  r
FT                   /mod_base= cm
FT   modified_base   26
FT                   /*tag=  s
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   27
FT                   /*tag=  t
FT                   /mod_base= um
FT   modified_base   28
FT                   /*tag=  u
FT                   /mod_base= gm
FT   modified_base   29
FT                   /*tag=  v
FT                   /mod_base= cm
FT   modified_base   30
FT                   /*tag=  w
FT                   /mod_base= cm
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-90Mb
CC   which contains a sequence which is complementary to at least two non-
CC   contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides complementary
CC   to a portion of the HBV RNA, may be used for inhibiting HBV replication
CC   in a cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 8 A; 5 C; 9 G; 2 T; 6 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity  100.0%;  Pred. No. 2.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||||||||||||
Db         11 TAAGGGTCGAUGUCCAUGCC 30


RESULT 4
AAT72620
ID   AAT72620 standard; DNA; 30 BP.
XX
AC   AAT72620;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-91b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-91b which
CC   contains a sequence which is complementary to at least two non-contiguous
CC   regions of a hepatitis B virus (HBV) nucleic acid. The antisense
CC   oligonucleotide may be used to detect the presence of HBV in a sample.
CC   The antisense oligonucleotide, and oligonucleotides complementary to a
CC   portion of the HBV RNA, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 7 A; 6 C; 11 G; 6 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   85.0%;  Pred. No. 2.6;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db         11 TAAGGGTCGATGTCCATGCC 30


RESULT 5
AAT72618
ID   AAT72618 standard; DNA; 30 BP.
XX
AC   AAT72618;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-90b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-90b which
CC   contains a sequence which is complementary to at least two non-contiguous
CC   regions of a hepatitis B virus (HBV) nucleic acid. The antisense
CC   oligonucleotide may be used to detect the presence of HBV in a sample.
CC   The antisense oligonucleotide, and oligonucleotides complementary to a
CC   portion of the HBV RNA, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 8 A; 5 C; 9 G; 8 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   85.0%;  Pred. No. 2.6;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db         11 TAAGGGTCGATGTCCATGCC 30


RESULT 6
AAT72621
ID   AAT72621 standard; DNA; 30 BP.
XX
AC   AAT72621;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-91Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        21. .30
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   21
FT                   /*tag=  c
FT                   /mod_base= um
FT   modified_base   22
FT                   /*tag=  d
FT                   /mod_base= gm
FT   modified_base   23
FT                   /*tag=  e
FT                   /mod_base= um
FT   modified_base   24
FT                   /*tag=  f
FT                   /mod_base= cm
FT   modified_base   25
FT                   /*tag=  g
FT                   /mod_base= cm
FT   modified_base   26
FT                   /*tag=  h
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   27
FT                   /*tag=  i
FT                   /mod_base= um
FT   modified_base   28
FT                   /*tag=  j
FT                   /mod_base= gm
FT   modified_base   29
FT                   /*tag=  k
FT                   /mod_base= cm
FT   modified_base   30
FT                   /*tag=  l
FT                   /mod_base= cm
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-91Mb
CC   which contains a sequence which is complementary to at least two non-
CC   contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides complementary
CC   to a portion of the HBV RNA, may be used for inhibiting HBV replication
CC   in a cell or for the treatment of HBV infection
```

```
XX
SQ   Sequence 30 BP; 7 A; 6 C; 11 G; 3 T; 3 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity  100.0%;  Pred. No. 2.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||||||||||||
Db         11 TAAGGGTCGAUGUCCAUGCC 30


RESULT 7
ADC64743/c
ID   ADC64743 standard; RNA; 31 BP.
XX
AC   ADC64743;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Hepatitis B virus DNA polymerase related RNA oligonucleotide.
XX
KW   screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
XX
OS   Synthetic.
OS   Hepatitis B virus.
XX
PN   KR2002007891-A.
XX
PD   29-JAN-2002.
XX
PF   19-JUL-2000; 2000KR-00041420.
XX
PR   19-JUL-2000; 2000KR-00041420.
XX
PA   (MOGA-) MOGAM BIOTECHNOLOGY INST.
PA   (VIRO-) VIROGEN CO LTD.
XX
PI   Ji HJ,  Jung SI,  Kim YC,  Min MG,  Ryu WS,  Yoon GS;
XX
DR   WPI; 2003-309015/30.
XX
PT   Screening of antiviral agents by protein-priming activity of hepatitis B
PT   virus DNA polymerase.
XX
PS   Disclosure; Page 12; 13pp; Korean.
XX
CC   The present invention describes a method of screening for an antiviral
CC   agent by the protein-priming activity of hepatitis B virus (HBV) DNA
CC   polymerase. Also described is developing an antiviral agent with a high
CC   selectivity to HBV which can be used for high-throughput screening. The
CC   present sequence represents an RNA oligonucleotide which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 31 BP; 10 A; 6 C; 8 G; 0 T; 7 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 10;  Length 31;
  Best Local Similarity   85.0%;  Pred. No. 2.6;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db         22 TAAGGGTCGATGTCCATGCC 3


RESULT 8
AAD27422/c
ID   AAD27422 standard; DNA; 639 BP.
XX
AC   AAD27422;
XX
DT   18-APR-2002  (first entry)
XX
DE   Hepatitis B virus (HBV) core antigen (HBcAg) encoding DNA #1.
XX
KW   Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
KW   prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KW   hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW   tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KW   dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KW   food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
KW   antiprotozoal; ds.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag=  a
FT                   /product= "HBcAg"
XX
PN   WO200198333-A2.
XX
PD   27-DEC-2001.
XX
PF   22-JUN-2001; 2001WO-GB002817.
XX
PR   22-JUN-2000; 2000GB-00015308.
PR   06-OCT-2000; 2000GB-00024544.
XX
PA   (CELL-) CELLTECH PHARM LTD.
XX
PI   Page M,  Li J,  Pumpens P;
XX
DR   WPI; 2002-098223/13.
DR   P-PSDB; AAE17018.
XX
PT   New proteins comprising a modified hepatitis B core antigen, useful as a
PT   vaccine in prophylactic or therapeutic vaccination of the human or animal
PT   body, particularly against hepatitis B virus infection.
XX
PS   Disclosure; Page 38-39; 40pp; English.
XX
CC   The invention relates to modified proteins comprising hepatitis B virus
CC   (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC   repeats has been deleted and the protein comprising the C-terminal
CC   cysteine of HBcAg. The deleted region may be replaced by an epitope from
CC   a protein other than HBcAg, in which case the HBcAg acts as a carrier to
CC   present the epitope to the immune system. This chimeric protein or its
CC   nucleic acid is useful as a vaccine or in a method of prophylactic or
CC   therapeutic vaccination of the human or animal body, particularly against
CC   HBV. The nucleic acid encoding the protein may be used in gene therapy or
CC   DNA vaccination protocols. The chimeric protein or its nucleic acid may
CC   also be used as the basis of a prophylactic vaccine against a range of
CC   diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
CC   influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
CC   immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
CC   tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
CC   diarrhoea, meningitis or gonorrhea. The present sequence is a DNA
CC   encoding Hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 6;  Length 639;
  Best Local Similarity   85.0%;  Pred. No. 3.3;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        104 TAAGGGTCGATGTCCATGCC 85


RESULT 9
AAD31509/c
ID   AAD31509 standard; DNA; 639 BP.
XX
AC   AAD31509;
```

```
XX
DT   18-JUN-2002  (first entry)
XX
DE   Hepatitis B virus core antigen (HBcAg) encoding DNA.
XX
KW   Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
KW   therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
KW   polio; herpes; rabies; AIDS; foot-and-mouth disease; ds.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag=  a
FT                   /product= "HBc protein"
FT   sig_peptide     1. .87
FT                   /*tag=  b
FT   mat_peptide     88. .636
FT                   /*tag=  c
FT                   /product= "Mature HBc protein"
XX
PN   WO200177158-A1.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-GB001607.
XX
PR   07-APR-2000; 2000EP-00107118.
XX
PA   (MEDE-) MEDEVA EURO LTD.
XX
PI   Gehin A,  Gilbert R,  Stuart D,  Rowlands D;
XX
DR   WPI; 2002-239995/29.
DR   P-PSDB; AAE19793.
XX
PT   Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for the
PT   prophylactic or therapeutic treatment of humans or animals against e.g.
PT   HB virus, viral hepatitis, hepatitis C virus, influenza, or foot-and-
PT   mouth disease.
XX
PS   Disclosure; Page 23-24; 27pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) core antigen
CC   (HBcAg) fusion proteins and polynucleotides encoding such proteins.
CC   Sequences of the invention are useful in methods of prophylactic or
CC   therapeutic vaccination or to manufacture medicaments for prophylactic or
CC   therapeutic vaccination of the human or animal body against HBV, e.g.
CC   against viral hepatitis. They are also useful as a prophylactic vaccine
CC   against e.g. hepatitis c virus, influenza, polio, herpes, rabies,
CC   acquired immune deficiency syndrome (AIDS) or foot-and-mouth disease. The
CC   present sequence is a DNA encoding hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 6;  Length 639;
  Best Local Similarity   85.0%;  Pred. No. 3.3;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        104 TAAGGGTCGATGTCCATGCC 85


RESULT 10
AAA71734/c
ID   AAA71734 standard; cDNA; 663 BP.
XX
AC   AAA71734;
XX
DT   06-AUG-2003  (revised)
DT   08-JAN-2001  (first entry)
XX
DE   HBV fusion protein comprising LHB and RGD encoding cDNA.
XX
KW   Fusion protein; protein coat; virus-specific packaging signal; psi;
KW   virus protein; cell permeability; cell-specific binding site; LHB;
KW   large surface protein; core antigen; gene therapy; ss.
XX
OS   Hepatitis B virus.
OS   Synthetic.
XX
PN   WO200046376-A2.
XX
PD   10-AUG-2000.
XX
PF   04-FEB-2000; 2000WO-DE000363.
XX
PR   05-FEB-1999;   99DE-01004800.
XX
PA   (HILD/) HILDT E.
XX
PI   Hildt E,  Hofschneider P;
XX
DR   WPI; 2000-514959/46.
DR   P-PSDB; AAB10596.
XX
PT   Particle for cell-specific gene delivery, useful in gene therapy,
PT   comprises nucleic acid in protein coat that includes a fusion protein of
PT   viral protein, permeability peptide and cell-binding site.
XX
PS   Claim 16a; Fig 1; 34pp; German.
XX
CC   This invention describes a novel particle (A), comprising a protein coat
CC   with a fusion protein (FP), and, inside the coat, a nucleic acid (I)
CC   including the sequence for a virus-specific packaging signal (psi) and a
CC   structural gene. FP contains a virus protein (VP), a peptide (P) that
CC   mediates cell permeability and a heterologous cell-specific binding site
CC   (RGD). The invention also describes (1) producing (A) in which FP
CC   contains an LHBs (large surface protein of hepatitis B virus (HBV)) and
CC   RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),
CC   (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector
CC   containing the DNA of (d). The products of the invention are used in gene
CC   therapy of cells and tissues, in vivo or ex vivo. This sequence encodes a
CC   fusion protein which is described in the method of the invention.
CC   (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 663 BP; 154 A; 169 C; 152 G; 188 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 3;  Length 663;
  Best Local Similarity   85.0%;  Pred. No. 3.3;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        107 TAAGGGTCGATGTCCATGCC 88


RESULT 11
ADO07220/c
ID   ADO07220 standard; DNA; 669 BP.
XX
AC   ADO07220;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Hepatitis B virus core antigen DNA.
XX
KW   HBcAg; immunomodulator; vaccine; gene; ss.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             10. .669
FT                   /*tag=  a
FT                   /product= "HBcAg"
```

```
FT                   /partial
FT                   /note= "No start codon"
XX
PN   WO2004035007-A2.
XX
PD   29-APR-2004.
XX
PF   17-OCT-2003; 2003WO-US033178.
XX
PR   17-OCT-2002; 2002US-0419279P.
XX
PA   (ORAG-) ORAGEN CORP.
XX
PI   Michaels F;
XX
DR   WPI; 2004-348329/32.
DR   P-PSDB; ADO07221.
XX
PT   Modulating a systemic immune response to a peptide in a mammal comprises
PT   transmucosally administering a macromolecular aggregate of the peptide.
XX
PS   Disclosure; SEQ ID NO 1; 81pp; English.
XX
CC   The present sequence is the DNA sequence of the hepatitis B virus core
CC   antigen (HBcAg) gene from HBV serotype ayw. A peptide comprising a HBV
CC   protein can be used in claimed methods of the invention for modulating an
CC   immune response in a mammal. A method of inducing a systemic immune
CC   response to a peptide in a mammal comprises transmucosally administering
CC   to the mammal a macromolecular aggregate of the peptide. The
CC   macromolecular aggregate comprises at least 10 peptide subunits, may have
CC   a molecular weight of over 1,000 kDa, and is preferably at least 5 nm in
CC   diameter. It is resistant to digestive degradation, being stabilised in
CC   aggregate form by chemical treatment and/or by recombinant protein
CC   engineering of the peptide. The peptide preferably comprises a HBV
CC   protein selected from HBV surface protein, nucleocapsid protein or
CC   envelope protein. Transmucosal administration to a mammal of a
CC   macromolecular aggregate of a HBV surface protein engenders a systemic
CC   immune response in the mammal. A method of suppressing an immune response
CC   in a mammal involves transmucosally administering a monomolecular peptide
CC   that is resistant to digestive degradation and which may be stabilised by
CC   chemical treatment or protein engineering, and which may be derived from
CC   a HBV protein. A monomolecular peptide is useful for the induction of
CC   oral tolerance when induction of systemic immunity is undesirable, e.g.
CC   in cases of chronic infections.
XX
SQ   Sequence 669 BP; 155 A; 170 C; 148 G; 196 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 12;  Length 669;
  Best Local Similarity   85.0%;  Pred. No. 3.3;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        134 TAAGGGTCGATGTCCATGCC 115


RESULT 12
AAV82691/c
ID   AAV82691 standard; DNA; 1334 BP.
XX
AC   AAV82691;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Fulminant hepatitis B virus genotype D variant FHBV12 sequence.
XX
KW   Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW   HBV-related disease; ss.
XX
OS   Hepatitis B virus.
XX
PN   WO9845421-A2.
XX
PD   15-OCT-1998.
XX
PF   08-APR-1998;   98WO-EP002048.
XX
PR   09-APR-1997;   97GB-00007221.
XX
PA   (UNIU ) UNIV GLASGOW.
XX
PI   Carman B;
XX
DR   WPI; 1999-009329/01.
XX
PT   New hepatitis B virus nucleic acid with combination of specific mutations
PT   - useful for, e.g. detection of binding interactions between host or
PT   viral proteins and HBV nucleic.
XX
PS   Disclosure; Fig 5; 85pp; English.
XX
CC   The present sequence represents part of the genome of a fulminant
CC   Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC   The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC   a mutation (i.e. alteration from the normal nucleotide in any of the
CC   genotypes A to F) in at least two of the enhancer I region, the negative
CC   regulatory element region, the enhancer II/ core upstream regulatory
CC   sequence/ basal core promoter region, or a mutation which leads to an X-
CC   peptide amino acid change to Cys or Met. The HBV variants of the
CC   invention are used to detect binding interactions between host or viral
CC   proteins and HBV nucleic acid. Probes that hybridise to any of the
CC   specified mutated regions are used to detect HBV-related disease,
CC   especially fulminant infection, but also severe chronic infection or
CC   serologically unusual forms of disease. Combinations of the specified
CC   mutations are associated with fulminant infections, probably because they
CC   reduce the ability to bind inhibitory proteins in the host cell
XX
SQ   Sequence 1334 BP; 288 A; 363 C; 311 G; 372 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 1334;
  Best Local Similarity   85.0%;  Pred. No. 3.5;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        806 TAAGGGTCGATGTCCATGCC 787


RESULT 13
AAV82688/c
ID   AAV82688 standard; DNA; 1395 BP.
XX
AC   AAV82688;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Fulminant hepatitis B virus genotype D variant FHBV5 sequence.
XX
KW   Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW   HBV-related disease; ss.
XX
OS   Hepatitis B virus.
XX
PN   WO9845421-A2.
XX
PD   15-OCT-1998.
XX
PF   08-APR-1998;   98WO-EP002048.
XX
PR   09-APR-1997;   97GB-00007221.
XX
PA   (UNIU ) UNIV GLASGOW.
XX
PI   Carman B;
XX
DR   WPI; 1999-009329/01.
```

```
XX
PT   New hepatitis B virus nucleic acid with combination of specific mutations
PT   - useful for, e.g. detection of binding interactions between host or
PT   viral proteins and HBV nucleic.
XX
PS   Disclosure; Fig 5; 85pp; English.
XX
CC   The present sequence represents part of the genome of a fulminant
CC   Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC   The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC   a mutation (i.e. alteration from the normal nucleotide in any of the
CC   genotypes A to F) in at least two of the enhancer I region, the negative
CC   regulatory element region, the enhancer II/ core upstream regulatory
CC   sequence/ basal core promoter region, or a mutation which leads to an X-
CC   peptide amino acid change to Cys or Met. The HBV variants of the
CC   invention are used to detect binding interactions between host or viral
CC   proteins and HBV nucleic acid. Probes that hybridise to any of the
CC   specified mutated regions are used to detect HBV-related disease,
CC   especially fulminant infection, but also severe chronic infection or
CC   serologically unusual forms of disease. Combinations of the specified
CC   mutations are associated with fulminant infections, probably because they
CC   reduce the ability to bind inhibitory proteins in the host cell
XX
SQ   Sequence 1395 BP; 277 A; 387 C; 331 G; 398 T; 0 U; 2 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 1395;
  Best Local Similarity   85.0%;  Pred. No. 3.5;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        917 TAAGGGTCGATGTCCATGCC 898


RESULT 14
AAV82687/c
ID   AAV82687 standard; DNA; 1400 BP.
XX
AC   AAV82687;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Fulminant hepatitis B virus genotype D variant FHBV4 sequence.
XX
KW   Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW   HBV-related disease; ss.
XX
OS   Hepatitis B virus.
XX
PN   WO9845421-A2.
XX
PD   15-OCT-1998.
XX
PF   08-APR-1998;   98WO-EP002048.
XX
PR   09-APR-1997;   97GB-00007221.
XX
PA   (UNIU ) UNIV GLASGOW.
XX
PI   Carman B;
XX
DR   WPI; 1999-009329/01.
XX
PT   New hepatitis B virus nucleic acid with combination of specific mutations
PT   - useful for, e.g. detection of binding interactions between host or
PT   viral proteins and HBV nucleic.
XX
PS   Disclosure; Fig 5; 85pp; English.
XX
CC   The present sequence represents part of the genome of a fulminant
CC   Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC   The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC   a mutation (i.e. alteration from the normal nucleotide in any of the
CC   genotypes A to F) in at least two of the enhancer I region, the negative
CC   regulatory element region, the enhancer II/ core upstream regulatory
CC   sequence/ basal core promoter region, or a mutation which leads to an X-
CC   peptide amino acid change to Cys or Met. The HBV variants of the
CC   invention are used to detect binding interactions between host or viral
CC   proteins and HBV nucleic acid. Probes that hybridise to any of the
CC   specified mutated regions are used to detect HBV-related disease,
CC   especially fulminant infection, but also severe chronic infection or
CC   serologically unusual forms of disease. Combinations of the specified
CC   mutations are associated with fulminant infections, probably because they
CC   reduce the ability to bind inhibitory proteins in the host cell
XX
SQ   Sequence 1400 BP; 287 A; 388 C; 332 G; 393 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 1400;
  Best Local Similarity   85.0%;  Pred. No. 3.5;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        917 TAAGGGTCGATGTCCATGCC 898


RESULT 15
AAV82692/c
ID   AAV82692 standard; DNA; 1445 BP.
XX
AC   AAV82692;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Fulminant hepatitis B virus genotype D variant FHBV13 sequence.
XX
KW   Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW   HBV-related disease; ss.
XX
OS   Hepatitis B virus.
XX
PN   WO9845421-A2.
XX
PD   15-OCT-1998.
XX
PF   08-APR-1998;   98WO-EP002048.
XX
PR   09-APR-1997;   97GB-00007221.
XX
PA   (UNIU ) UNIV GLASGOW.
XX
PI   Carman B;
XX
DR   WPI; 1999-009329/01.
XX
PT   New hepatitis B virus nucleic acid with combination of specific mutations
PT   - useful for, e.g. detection of binding interactions between host or
PT   viral proteins and HBV nucleic.
XX
PS   Disclosure; Fig 5; 85pp; English.
XX
CC   The present sequence represents part of the genome of a fulminant
CC   Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC   The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC   a mutation (i.e. alteration from the normal nucleotide in any of the
CC   genotypes A to F) in at least two of the enhancer I region, the negative
CC   regulatory element region, the enhancer II/ core upstream regulatory
CC   sequence/ basal core promoter region, or a mutation which leads to an X-
CC   peptide amino acid change to Cys or Met. The HBV variants of the
CC   invention are used to detect binding interactions between host or viral
CC   proteins and HBV nucleic acid. Probes that hybridise to any of the
CC   specified mutated regions are used to detect HBV-related disease,
CC   especially fulminant infection, but also severe chronic infection or
CC   serologically unusual forms of disease. Combinations of the specified
CC   mutations are associated with fulminant infections, probably because they
CC   reduce the ability to bind inhibitory proteins in the host cell
```

XX
SQ Sequence 1445 BP; 297 A; 406 C; 338 G; 404 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 1445;
Best Local Similarity 85.0%; Pred. No. 3.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||||||:|:|||:|||
Db        917 TAAGGGTCGATGTCCATGCC 898
```

Search completed: March 17, 2005, 06:48:43
Job time : 172.333 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-62
Perfect score: 20
Sequence: 1 taagggtcgauguccaugcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	19	95.0	686	4	BI954819	BI954819 HVSMEm001
	2	17.4	87.0	576	8	CC145068	CC145068 ZMMBBb000
	3	17.4	87.0	763	9	CL248902	CL248902 ZMMBBb059
c	4	17	85.0	260	7	CN589115	CN589115 TTE000135
	5	17	85.0	451	5	BP643309	BP643309 BP643309
c	6	17	85.0	729	4	BJ705222	BJ705222 BJ705222
	7	16.8	84.0	413	5	BP569691	BP569691 BP569691
	8	16.8	84.0	431	9	CL260820	CL260820 ZMMBBb061
	9	16.8	84.0	457	5	BP527907	BP527907 BP527907
	10	16.8	84.0	509	9	TA230D01P	AL481016 T. brucei
	11	16.8	84.0	733	1	AI635930	AI635930 tz82c11.x
c	12	16.8	84.0	773	8	BZ573222	BZ573222 msh2_3006
c	13	16.8	84.0	1574	9	AG476230	AG476230 Mus muscu
	14	16.4	82.0	272	2	BE922196	BE922196 EST425953
c	15	16.4	82.0	306	9	AG200800	AG200800 Pan trogl
	16	16.4	82.0	356	9	CC836806	CC836806 ZMMBBc048
c	17	16.4	82.0	417	7	H70178	H70178 yr94b08.r1
	18	16.4	82.0	499	5	BQ116265	BQ116265 EST601828
c	19	16.4	82.0	554	8	AQ433745	AQ433745 HS_5063_A
c	20	16.4	82.0	612	9	CL194160	CL194160 104_418_1
	21	16.4	82.0	622	5	BQ112668	BQ112668 EST598244
	22	16.4	82.0	633	9	CG461236	CG461236 PUFXP93TD
	23	16.4	82.0	678	9	CG761335	CG761335 ZMMBBb033
c	24	16.4	82.0	684	9	CL369255	CL369255 RPCI44_29
	25	16.4	82.0	714	9	AG128210	AG128210 Pan trogl
	26	16.4	82.0	772	8	BH257005	BH257005 CH230-283
c	27	16.4	82.0	848	7	CO099982	CO099982 GR__Ea25C
	28	16.4	82.0	851	4	BJ571152	BJ571152 BJ571152
	29	16.4	82.0	924	9	CG289400	CG289400 OGXBN55TV
	30	16.4	82.0	984	9	CG875596	CG875596 ZMMBBc028
	31	16.4	82.0	1010	9	CC622611	CC622611 OGULF56TV
c	32	16.4	82.0	1046	9	CL055525	CL055525 CH216-81E
	33	16.4	82.0	1669	3	HSM800819	AL110272 Homo sapi
	34	16.4	82.0	1669	3	HSM802167	AL137702 Homo sapi
	35	16	80.0	168	4	BI052244	BI052244 PM2-GN037
	36	16	80.0	168	4	BI052309	BI052309 PM2-GN037
	37	16	80.0	532	5	BX433931	BX433931 BX433931
c	38	16	80.0	666	9	CG106546	CG106546 PUIJJ59TB
c	39	16	80.0	698	8	BZ413874	BZ413874 if18d10.g
	40	16	80.0	740	9	CG235822	CG235822 OGWKA72TV
c	41	16	80.0	750	9	CG235810	CG235810 OGWKA72TH
c	42	16	80.0	810	9	CC716058	CC716058 OGTAI38TC
c	43	16	80.0	836	9	CG257272	CG257272 OG1CX72TH
	44	16	80.0	963	9	CG302315	CG302315 OG2BM75TV
	45	15.8	79.0	68	8	BH631190	BH631190 1007074G0

ALIGNMENTS

RESULT 1
BI954819/c
LOCUS BI954819 686 bp mRNA linear EST 19-OCT-2001
DEFINITION HVSMEm0019O17f Hordeum vulgare green seedling EST library HVcDNA0014 (Blumeria infected) Hordeum vulgare subsp. vulgare cDNA clone HVSMEm0019O17f, mRNA sequence.
ACCESSION BI954819
VERSION BI954819.1 GI:16300646
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 686)
AUTHORS Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Chin,A., Begum,D., Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Oates,R. and Main,D.
TITLE Development of a genetically and physically anchored EST resource for barley genomics: Blumeria infected Morex (compatible) seedling cDNA library
JOURNAL Unpublished (2001)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hq bases = 417
Seq primer: AATTAACCCTCACTAAAGGG
High quality sequence start: 16
High quality sequence stop: 531.
FEATURES Location/Qualifiers
source 1. .686
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSMEm0019O17f"
/tissue_type="green seedling leaf"
/lab_host="TJC121"
/clone_lib="Hordeum vulgare green seedling EST library HVcDNA0014 (Blumeria infected)"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Morex (mla) plants were greenhouse grown in the R

Wise lab at Iowa State University, Ames, IA; 7 day old green seedlings were infected with isolate 5874 of Blumeria graminis f. sp. hordei, and leaves were harvested 24, 48 and 72 hr post-inoculation and snap frozen (Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Chin). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

ORIGIN

```
  Query Match             95.0%;  Score 19;  DB 4;  Length 686;
  Best Local Similarity   84.2%;  Pred. No. 63;
  Matches   16;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AAGGGTCGAUGUCCAUGCC 20
              |||||||||:|:|||:|||
Db        136 AAGGGTCGATGTCCATGCC 118
```

RESULT 2
CC145068

```
LOCUS       CC145068                576 bp    DNA     linear   GSS 24-JUN-2003
DEFINITION  ZMMBBb0002P18.r ZMMBBb Zea mays genomic clone ZMMBBb0002P18 3',
            genomic survey sequence.
ACCESSION   CC145068
VERSION     CC145068.1  GI:30090261
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 576)
  AUTHORS   Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
            and Wing,R.
  TITLE     Sequencing of the maize genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: T7
            BACKWARD: M13r
            Plate: 0002  row: P  column: 18
            Seq primer: M13r
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .576
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0002P18"
                     /lab_host="DH10B"
                     /clone_lib="ZMMBBb"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Zea mays L. ssp. mays"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 8;  Length 576;
  Best Local Similarity   84.2%;  Pred. No. 4.2e+02;
  Matches   16;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAGGGTCGAUGUCCAUGCC 20
              |||||||||:| |||:|||
Db        139 AAGGGTCGATGCCCATGCC 157
```

RESULT 3
CL248902

```
LOCUS       CL248902                763 bp    DNA     linear   GSS 22-JAN-2004
DEFINITION  ZMMBBb0597N17r ZMMBBb (HindIII) Zea mays genomic clone
            ZMMBBb0597N17 3', genomic survey sequence.
ACCESSION   CL248902
VERSION     CL248902.1  GI:41105456
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 763)
  AUTHORS   Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
            Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
  TITLE     Sequencing of the maize genome at PGIR (2003c)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Bharti,A.K.
            Dr.Joachim Messing's lab
            The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
            University
            190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            Tel: 732 445 3801
            Fax: 732 445 5735
            Email: bharti@waksman.rutgers.edu
            Seq primer: SP6
            Class: BAC ends
            High quality sequence start: 405.
FEATURES             Location/Qualifiers
     source          1. .763
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0597N17"
                     /lab_host="E. coli DH10B"
                     /clone_lib="ZMMBBb (HindIII)"
                     /note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 763;
  Best Local Similarity   78.9%;  Pred. No. 4.4e+02;
  Matches   15;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAGGGTCGAUGUCCAUGCC 20
              |||||||||:|:| |:|||
Db        339 AAGGGTCGATGTCTATGCC 357
```

RESULT 4
CN589115/c

```
LOCUS       CN589115                260 bp    mRNA    linear   EST 31-AUG-2004
DEFINITION  TTE00013587 Normalized large Tetrahymena thermophila cDNA, mRNA
            sequence.
ACCESSION   CN589115
```

```
VERSION     CN589115.1  GI:47040917
KEYWORDS    EST.
SOURCE      Tetrahymena thermophila
  ORGANISM  Tetrahymena thermophila
            Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
            Hymenostomatida; Tetrahymenina; Tetrahymenidae; Tetrahymena.
REFERENCE   1  (bases 1 to 260)
  AUTHORS   Garg,J., Pearlman,R.E. and Carlton,J.
  TITLE     PEPdbPub (http://amoebidia.bcm.umontreal.ca/public/pepdb/agrm.php)
            Tetrahymena thermophila (TIGR)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: PEPdb
            Departement de Biochimie, Universite de Montreal
            Email: pepdb-curator@bch.umontreal.ca
            Plate: 1398.
FEATURES             Location/Qualifiers
     source          1. .260
                     /organism="Tetrahymena thermophila"
                     /mol_type="mRNA"
                     /db_xref="taxon:5911"
                     /clone_lib="Normalized large"
ORIGIN

  Query Match            85.0%;  Score 17;  DB 7;  Length 260;
  Best Local Similarity  82.4%;  Pred. No. 6e+02;
  Matches   14;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GGGTCGAUGUCCAUGCC 20
              |||||||:|:|||:|||
Db        237 GGGTCGATGTCCATGCC 221


RESULT 5
BP643309
LOCUS       BP643309                451 bp    mRNA    linear   EST 27-JUN-2004
DEFINITION  BP643309 RAFL19 Arabidopsis thaliana cDNA clone RAFL19-61-I08 3',
            mRNA sequence.
ACCESSION   BP643309
VERSION     BP643309.1  GI:49294779
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
            Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
            Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
  TITLE     Functional annotation of a full-length Arabidopsis cDNA collection
  JOURNAL   Science 296 (5565), 141-145 (2002)
  MEDLINE   21932900
   PUBMED   11910074
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            reversed clone; Please visit our web site
            (http://pfgweb.gsc.riken.go.jp/) for further details.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL19-61-I08"
                     /tissue_type="mixture of silique and flower"
                     /lab_host="DH10B"
                     /clone_lib="RAFL19"
                     /note="Site_1: BamHI; Site_2: SalI; Subtraction Library"
ORIGIN

  Query Match            85.0%;  Score 17;  DB 5;  Length 451;
  Best Local Similarity  82.4%;  Pred. No. 6.5e+02;
  Matches   14;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 AGGGTCGAUGUCCAUGC 19
              ||||||||:|:|||:||
Db        431 AGGGTCGATGTCCATGC 447


RESULT 6
BJ705222/c
LOCUS       BJ705222                729 bp    mRNA    linear   EST 08-MAR-2004
DEFINITION  BJ705222 MF01FFA cDNA Oryzias latipes cDNA clone MF01FFA013a11 5',
            mRNA sequence.
ACCESSION   BJ705222
VERSION     BJ705222.1  GI:45246102
KEYWORDS    EST.
SOURCE      Oryzias latipes (Japanese medaka)
  ORGANISM  Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
  TITLE     Medaka EST Project in Takeda's lab
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Oryzias latipes"
                     /mol_type="mRNA"
                     /strain="Hd-rR"
                     /db_xref="taxon:8090"
                     /clone="MF01FFA013a11"
                     /sex="mixture of female and male"
                     /tissue_type="whole embryo"
                     /dev_stage="fry stage 40"
                     /clone_lib="MF01FFA cDNA"
ORIGIN

  Query Match            85.0%;  Score 17;  DB 4;  Length 729;
  Best Local Similarity  82.4%;  Pred. No. 7e+02;
  Matches   14;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AAGGGTCGAUGUCCAUG 18
              |||||||||:|:|||:|
Db        182 AAGGGTCGATGTCCATG 166


RESULT 7
BP569691
LOCUS       BP569691                413 bp    mRNA    linear   EST 20-JUN-2004
DEFINITION  BP569691 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-68-M24 3',
            mRNA sequence.
ACCESSION   BP569691
VERSION     BP569691.1  GI:48985457
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
```

```
            Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
            Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
  TITLE     Functional annotation of a full-length Arabidopsis cDNA collection
  JOURNAL   Science 296 (5565), 141-145 (2002)
  MEDLINE   21932900
   PUBMED   11910074
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            reversed clone; Please visit our web site
            (http://pfgweb.gsc.riken.go.jp/) for further details.
FEATURES             Location/Qualifiers
     source          1. .413
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL14-68-M24"
                     /tissue_type="root"
                     /lab_host="DH10B"
                     /clone_lib="RAFL14"
                     /note="Site_1: BamHI; Site_2: SalI"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 5;  Length 413;
  Best Local Similarity   75.0%;  Pred. No. 8.2e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||| ||:|:|||:| |
Db        383 TAAGGGTTGATGTCCATGAC 402


RESULT 8
CL260820
LOCUS       CL260820                431 bp    DNA     linear   GSS 02-FEB-2004
DEFINITION  ZMMBBb0619B24r ZMMBBb (HindIII) Zea mays genomic clone
            ZMMBBb0619B24 3', genomic survey sequence.
ACCESSION   CL260820
VERSION     CL260820.1  GI:41624959
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
            Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
  TITLE     Sequencing of the maize genome at PGIR (2003c)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Bharti,A.K.
            Dr.Joachim Messing's lab
            The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
            University
            190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            Tel: 732 445 3801
            Fax: 732 445 5735
            Email: bharti@waksman.rutgers.edu
            Seq primer: SP6
            Class: BAC ends
            High quality sequence start: 116.
FEATURES             Location/Qualifiers
     source          1. .431
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0619B24"
                     /lab_host="E. coli DH10B"
                     /clone_lib="ZMMBBb (HindIII)"
                     /note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 431;
  Best Local Similarity   85.0%;  Pred. No. 8.3e+02;
  Matches   17;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              |||||||||:| ||| |||
Db        328 TAAGGGTCGATGCCCAGGCC 347


RESULT 9
BP527907
LOCUS       BP527907                457 bp    mRNA    linear   EST 28-SEP-2004
DEFINITION  BP527907 MAT001 Nicotiana tabacum cDNA clone BY12728, mRNA
            sequence.
ACCESSION   BP527907
VERSION     BP527907.1  GI:52831634
KEYWORDS    EST.
SOURCE      Nicotiana tabacum (common tobacco)
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Matsuoka,K., Tashiro,G., Horiguchi,T., Demura,T. and Fukuda,H.
  TITLE     Profiling growth-phase dependent gene expression of tobacco BY-2
            cells by comprehensive microarray analysis
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ken Matsuoka
            Morphogenesis Research Group
            RIKEN Plant Science Center
            1-7-2 Suehirocho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9575
            Fax: 81-45-503-9573
            Email: by2@psc.riken.go.jp, URL:http://mrg.psc.riken.go.jp/strc/
            The cDNA library was constructed from mRNA isolated from lag (9 h),
            log (72 h) and stationary (7 days) old BY-2 cells.
            Seq primer: M13 forward.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /cultivar="Bright Yellow No.2"
                     /db_xref="taxon:4097"
                     /clone="BY12728"
                     /cell_line="BY-2"
                     /clone_lib="MAT001"
                     /note="Vector: pGEM-T easy; primer: M13 forward; mRNA
                     obtained from lag, log and stationary phase cells"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 5;  Length 457;
  Best Local Similarity   75.0%;  Pred. No. 8.3e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||| | ||||:|:|||:|||
Db        297 TAAAGTTCGATGTCCATGCC 316


RESULT 10
TA230D01P
LOCUS       TA230D01P               509 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 230d01, forward sequence,
            genomic survey sequence.
ACCESSION   AL481016
VERSION     AL481016.1  GI:11846785
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
```

```
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 509)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES              Location/Qualifiers
     source           1. .509
                      /organism="Trypanosoma brucei"
                      /mol_type="genomic DNA"
                      /strain="TREU927"
                      /db_xref="taxon:5691"
                      /clone="230d01"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 509;
  Best Local Similarity   80.0%;  Pred. No. 8.5e+02;
  Matches   16;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||||||| ||:|:||| |||
Db        355 TAAGGGTTGATGTCCAGGCC 374


RESULT 11
AI635930
LOCUS       AI635930                733 bp    mRNA    linear   EST 16-DEC-1999
DEFINITION  tz82c11.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2295092 3'
            similar to gb:J03490 DIHYDROLIPOAMIDE DEHYDROGENASE PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   AI635930
VERSION     AI635930.1  GI:4687260
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 733)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11548-013
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1107   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 401.
FEATURES              Location/Qualifiers
     source           1. .733
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2295092"
                      /tissue_type="adenocarcinoma"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Pan1"
                      /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.72 kb. Life Technologies catalog #:
                      11548-013"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 1;  Length 733;
  Best Local Similarity   75.0%;  Pred. No. 8.9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              |||||||||:|: ||:| |
Db        557 TAAGGGTCGATGTGCATGAC 576


RESULT 12
BZ573222/c
LOCUS       BZ573222                773 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  msh2_3006.y2 msh Pseudomonas aeruginosa genomic clone msh2_3006,
            genomic survey sequence.
ACCESSION   BZ573222
VERSION     BZ573222.1  GI:27208283
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES              Location/Qualifiers
     source           1. .773
                      /organism="Pseudomonas aeruginosa"
                      /mol_type="genomic DNA"
                      /strain="MSH"
                      /db_xref="taxon:287"
                      /clone="msh2_3006"
                      /clone_lib="msh"
                      /note="Environmental isolate. Whole genomic shotgun
                      library."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 8;  Length 773;
  Best Local Similarity   80.0%;  Pred. No. 9e+02;
  Matches   16;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              ||| ||||||:|:||| |||
Db        294 TAATGGTCGATGTCCAGGCC 275


RESULT 13
AG476230/c
LOCUS       AG476230                1574 bp    DNA     linear   GSS 04-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-369C21.TJ, genomic survey
            sequence.
```

```
ACCESSION   AG476230
VERSION     AG476230.1  GI:48183460
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1574)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1574
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-369C21.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 1574;
  Best Local Similarity   75.0%;  Pred. No. 1e+03;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGTCGAUGUCCAUGCC 20
              || ||||||:|:|||: ||
Db        300 TATGGGTCGATGTCCATCCC 281


RESULT 14
BE922196
LOCUS       BE922196                272 bp    mRNA    linear   EST 07-MAR-2003
DEFINITION  EST425953 potato leaves and petioles Solanum tuberosum cDNA clone
            cSTB18G4 5' sequence, mRNA sequence.
ACCESSION   BE922196
VERSION     BE922196.1  GI:10448260
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 272)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
            Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R.,
            Ronning,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
  TITLE     Generation of ESTs from potato leaves and petioles
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/.
FEATURES             Location/Qualifiers
     source          1. .272
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="cSTB18G4"
                     /tissue_type="leaflets and petioles"
                     /dev_stage="8 weeks old plants"
                     /lab_host="SOLR"
                     /clone_lib="potato leaves and petioles"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; Tissue was supplied by Dr. Fry (Cornell University).
                     Leaflets and petioles were isolated from 8 week old
                     greenhouse grown plants. The plants were watered and
                     fertilized freely. The tissue was immediately frozen in
                     liquid nitrogen."
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 2;  Length 272;
  Best Local Similarity   77.8%;  Pred. No. 1.2e+03;
  Matches   14;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 AGGGTCGAUGUCCAUGCC 20
              ||| ||||:|:|||:|||
Db         29 AGGTTCGATGTCCATGCC 46


RESULT 15
AG200800/c
LOCUS       AG200800                306 bp    DNA     linear   GSS 06-MAR-2004
DEFINITION  Pan troglodytes DNA, clone: RP43-082P16.T7, genomic survey
            sequence.
ACCESSION   AG200800
VERSION     AG200800.1  GI:45232975
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     BAC end sequences of Library RP-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 306)
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of
            Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC);
            52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea
            (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/,
            Tel:82-42-866-7181, Fax:82-42-860-4409)
COMMENT     Clones are derived from the chimpanzee BAC library RP-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
            PRIMERS
                Sequencing: T7
            LIBRARY
                Vector     : pBACe3.6
                R.Site 1   : EcoRI
                R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .306
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
```

/db_xref="taxon:9598"
/clone="RP43-082P16.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RP-43 Chimpanzee Male BAC Library"
ORIGIN

Query Match 82.0%; Score 16.4; DB 9; Length 306;
Best Local Similarity 77.8%; Pred. No. 1.3e+03;
Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TAAGGGTCGAUGUCCAUG 18
              ||||||| ||:|:|||:|
Db         35 TAAGGGTAGATGTCCATG 18
```

Search completed: March 17, 2005, 11:07:49
Job time : 1389.27 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-63
Perfect score: 20
Sequence: 1 taagggucgauguccatgcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	6	AR027817	AR027817 Sequence
2	20	100.0	30	6	AR027842	AR027842 Sequence
3	20	100.0	30	6	AR027843	AR027843 Sequence
c 4	20	100.0	81	6	I92348	I92348 Sequence 9
c 5	20	100.0	174	14	S77749	S77749 preC {X/pre
c 6	20	100.0	253	14	AY329529	AY329529 Hepatitis
c 7	20	100.0	253	14	AY329562	AY329562 Hepatitis
c 8	20	100.0	253	14	AY329568	AY329568 Hepatitis
c 9	20	100.0	253	14	AY329573	AY329573 Hepatitis
c 10	20	100.0	253	14	AY329581	AY329581 Hepatitis
c 11	20	100.0	294	14	AF390000	AF390000 Hepatitis
c 12	20	100.0	333	14	HPBHBED	L12359 Hepatitis B
c 13	20	100.0	398	14	AB167603	AB167603 Hepatitis
c 14	20	100.0	398	14	AB167637	AB167637 Hepatitis
c 15	20	100.0	406	14	AB163815	AB163815 Hepatitis
c 16	20	100.0	406	14	AB163817	AB163817 Hepatitis
c 17	20	100.0	439	14	AY254503	AY254503 Hepatitis
c 18	20	100.0	488	14	AY274419	AY274419 Hepatitis
c 19	20	100.0	488	14	AY274420	AY274420 Hepatitis
c 20	20	100.0	488	14	AY274422	AY274422 Hepatitis
c 21	20	100.0	488	14	AY274427	AY274427 Hepatitis
c 22	20	100.0	488	14	AY274428	AY274428 Hepatitis
c 23	20	100.0	488	14	AY274429	AY274429 Hepatitis
c 24	20	100.0	488	14	AY274430	AY274430 Hepatitis
c 25	20	100.0	488	14	AY274431	AY274431 Hepatitis
c 26	20	100.0	488	14	AY274432	AY274432 Hepatitis
c 27	20	100.0	488	14	AY274433	AY274433 Hepatitis
c 28	20	100.0	488	14	AY274434	AY274434 Hepatitis
c 29	20	100.0	488	14	AY274436	AY274436 Hepatitis
c 30	20	100.0	493	14	S79556	S79556 X, preC [he
c 31	20	100.0	548	14	AY382500	AY382500 Hepatitis
c 32	20	100.0	548	14	AY382501	AY382501 Hepatitis
c 33	20	100.0	548	14	AY382502	AY382502 Hepatitis
c 34	20	100.0	548	14	AY382521	AY382521 Hepatitis
c 35	20	100.0	548	14	AY382522	AY382522 Hepatitis
c 36	20	100.0	548	14	AY382523	AY382523 Hepatitis
c 37	20	100.0	548	14	AY382524	AY382524 Hepatitis
c 38	20	100.0	548	14	AY382525	AY382525 Hepatitis
c 39	20	100.0	548	14	AY382526	AY382526 Hepatitis
c 40	20	100.0	548	14	AY382527	AY382527 Hepatitis
c 41	20	100.0	609	14	AF289954	AF289954 Hepatitis
c 42	20	100.0	609	14	AF289965	AF289965 Hepatitis
c 43	20	100.0	626	14	AY254500	AY254500 Hepatitis
c 44	20	100.0	639	6	AX278066	AX278066 Sequence
c 45	20	100.0	639	6	AX342485	AX342485 Sequence

ALIGNMENTS

RESULT 1
AR027817
LOCUS AR027817 20 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 15 from patent US 5856459.
ACCESSION AR027817
VERSION AR027817.1 GI:5938637
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and Mills,J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 15 05-JAN-1999;
FEATURES Location/Qualifiers
source 1. .20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 85.0%; Pred. No. 14;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAAGGGUCGAUGUCCATGCC 20
||||||:|||:|:|||||||
Db 1 TAAGGGTCGATGTCCATGCC 20

RESULT 2
AR027842
LOCUS AR027842 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 40 from patent US 5856459.
ACCESSION AR027842
VERSION AR027842.1 GI:5938662
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and

```
            Mills,J.S.
  TITLE     Oligonucleotides specific for hepatitis B virus
  JOURNAL   Patent: US 5856459-A 40 05-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 30;
  Best Local Similarity   85.0%;  Pred. No. 14;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db         11 TAAGGGTCGATGTCCATGCC 30


RESULT 3
AR027843
LOCUS       AR027843                 30 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 41 from patent US 5856459.
ACCESSION   AR027843
VERSION     AR027843.1  GI:5938663
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and
            Mills,J.S.
  TITLE     Oligonucleotides specific for hepatitis B virus
  JOURNAL   Patent: US 5856459-A 41 05-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 30;
  Best Local Similarity   85.0%;  Pred. No. 14;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db         11 TAAGGGTCGATGTCCATGCC 30


RESULT 4
I92348/c
LOCUS       I92348                   81 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 9 from patent US 5728518.
ACCESSION   I92348
VERSION     I92348.1  GI:3936818
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Carmichael,E.
  TITLE     Antiviral poly-and oligonucleotides
  JOURNAL   Patent: US 5728518-A 9 17-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .81
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 81;
  Best Local Similarity   85.0%;  Pred. No. 15;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db         71 TAAGGGTCGATGTCCATGCC 52


RESULT 5
S77749/c
LOCUS       S77749                   174 bp    DNA     linear   VRL 06-MAY-2003
DEFINITION  preC {X/preC region, deletion mutant} [hepatitis B virus HBV,
            host=human, serum, patient 5 isolate, Genomic DNA Mutant, 174 nt].
ACCESSION   S77749
VERSION     S77749.1  GI:999129
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 174)
  AUTHORS   Feitelson,M.A., Duan,L.X., Guo,J. and Blumberg,B.S.
  TITLE     X region deletion mutants associated with surface antigen-positive
            hepatitis B virus infections
  JOURNAL   Gastroenterology 108 (6), 1810-1819 (1995)
  MEDLINE   95285997
   PUBMED   7768387
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 165980] from the original journal article.
FEATURES             Location/Qualifiers
     source          1. .174
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
     gene            69. .>174
                     /gene="preC"
                     /note="no start codon found"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 174;
  Best Local Similarity   85.0%;  Pred. No. 16;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        172 TAAGGGTCGATGTCCATGCC 153


RESULT 6
AY329529/c
LOCUS       AY329529                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate A611252E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329529
VERSION     AY329529.1  GI:37625315
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="A611252E"
```

```
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95861.1"
                     /db_xref="GI:37625316"
                     /translation="STTDLEAYFKDCVFKDWEELGEETRLMIFVLGGCRHKLVCAPAS
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95862.1"
                     /db_xref="GI:37625317"
                     /translation="MQLFHLCLVISCTCPTFQASKLCLGWL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 7
AY329562/c
LOCUS       AY329562                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D273984E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329562
VERSION     AY329562.1  GI:37625413
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D273984E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95926.1"
                     /db_xref="GI:37625414"
                     /translation="STTDLEAYFKDCLFKDWEELGEEIRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95927.1"
                     /db_xref="GI:37625415"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 8
AY329568/c
LOCUS       AY329568                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D296668E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329568
VERSION     AY329568.1  GI:37625431
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D296668E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95938.1"
                     /db_xref="GI:37625432"
                     /translation="STTDLEAYFKDCLFKDWEELGEELRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95939.1"
                     /db_xref="GI:37625433"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 9
AY329573/c
LOCUS       AY329573                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D604917E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329573
VERSION     AY329573.1  GI:37625446
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
```

```
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D604917E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95948.1"
                     /db_xref="GI:37625447"
                     /translation="STTDLEAYFKDCLFKDWEELGEEIRLKVFVLGGCRHKLVCAPAS
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95949.1"
                     /db_xref="GI:37625448"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 10
AY329581/c
LOCUS       AY329581                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D639472E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329581
VERSION     AY329581.1  GI:37625470
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D639472E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95964.1"
                     /db_xref="GI:37625471"
                     /translation="STTDLEAYFKDCLFKDWEELGEELRLLIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95965.1"
                     /db_xref="GI:37625472"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        237 TAAGGGTCGATGTCCATGCC 218


RESULT 11
AF390000/c
LOCUS       AF390000                 294 bp    DNA     linear   VRL 06-MAR-2002
DEFINITION  Hepatitis B virus isolate D3 X protein gene, partial cds; and
            nonfunctional precore/core protein gene, partial sequence.
ACCESSION   AF390000
VERSION     AF390000.1  GI:16266099
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   Castro,L.D., Niel,C. and Gomes,S.A.
  TITLE     Low frequency of mutations in the core promoter and precore regions
            of hepatitis B virus in anti-HBe positive Brazilian carriers
  JOURNAL   BMC Microbiol. 1 (1), 10 (2001)
   PUBMED   11472634
REFERENCE   2  (bases 1 to 294)
  AUTHORS   De Castro,L., Niel,C. and Gomes,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-2001) Virology, FIOCRUZ, Av. Brasil 4365, Rio de
            Janeiro, RJ 21045-900, Brazil
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D3"
                     /db_xref="taxon:10407"
     CDS             <1. .119
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAL16752.1"
                     /db_xref="GI:16266100"
                     /translation="FKDWEELGDDSRLMIYVLGGCRHKLVCAPAPCNFFTSA"
     misc_feature    95. .>294
                     /note="nonfunctional precore/core protein due to mutation"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 294;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        198 TAAGGGTCGATGTCCATGCC 179
```

```
RESULT 12
HPBHBED/c
LOCUS       HPBHBED                  333 bp    DNA     linear   VRL 28-JUN-1996
DEFINITION  Hepatitis B virus precore and core protein gene, 5' end of cds.
ACCESSION   L12359
VERSION     L12359.1  GI:306267
KEYWORDS    HBcAg protein; HBeAg protein; core protein; nucleotide binding
            protein; precore protein.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (sites)
  AUTHORS   Tong,S.P., Li,J.S., Vitvitski,L. and Trepo,C.
  TITLE     Active hepatitis B virus replication in the presence of anti-HBe is
            associated with viral variants containing an inactive pre-C region
  JOURNAL   Virology 176 (2), 596-603 (1990)
  MEDLINE   90266476
   PUBMED   2345966
REFERENCE   2  (bases 1 to 333)
  AUTHORS   Li,J.S., Tong,S.P., Wen,Y.M., Vitvitski,L., Zhang,Q. and Trepo,C.
  TITLE     Hepatitis B virus genotype A rarely circulates as an HBe-minus
            mutant: possible contribution of a single nucleotide in the precore
            region
  JOURNAL   J. Virol. 67 (9), 5402-5410 (1993)
  MEDLINE   93353617
   PUBMED   8350403
COMMENT     Original source text: Hepatitis B virus DNA.
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
                     /note="genotype D; from French HBeAg- patient"
     CDS             1. .>333
                     /standard_name="HBeAg"
                     /pseudo
                     /codon_start=1
                     /product="precore protein"
     variation       82. .84
                     /note="HBe-abolishing"
     CDS             88. .>333
                     /standard_name="HBcAb"
                     /codon_start=1
                     /protein_id="AAB03102.1"
                     /db_xref="GI:306268"
                     /translation="MDIDPYKEFGATVELLTFLPSDFFPSVRDLLDTAAALYREALES
                     PEHCSHHHTALRQAILCWGELMTLATWVGANLDDPASR"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 333;
  Best Local Similarity   85.0%;  Pred. No. 17;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        104 TAAGGGTCGATGTCCATGCC 85


RESULT 13
AB167603/c
LOCUS       AB167603                 398 bp    DNA     linear   VRL 01-OCT-2004
DEFINITION  Hepatitis B virus gene for polyprotein, partial cds, clone: NEP75.
ACCESSION   AB167603
VERSION     AB167603.1  GI:53148166
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1
  AUTHORS   Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
  TITLE     A Case-control Study for Differences among Hepatitis B Virus
            Infections of Genotypes A (Subtypes Aa and Ae) and D
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 398)
  AUTHORS   Tanaka,Y. and Mizokami,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-2004) Yasuhito Tanaka, Nagoya City University
            Graduate School of Medical Sciences, Department of Clinical
            Molecular Informative Medicine; 1 Kawasumi, Mizuho-cho, Mizuho-ku,
            Nagoya, Aichi 467-8601, Japan (E-mail:ytanaka@med.nagoya-cu.ac.jp,
            Tel:81-52-853-8292, Fax:81-52-842-0021)
FEATURES             Location/Qualifiers
     source          1. .398
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
                     /clone="NEP75"
     CDS             184. .>398
                     /note="precore-core region"
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="BAD52175.1"
                     /db_xref="GI:53148167"
                     /translation="MQLFHLCLIISCTCPTFQASKLCLGWLWGMDIDPYKEFGATVEL
                     LSFLPSDFFPSVRDLLDTASALYREAL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 398;
  Best Local Similarity   85.0%;  Pred. No. 18;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        287 TAAGGGTCGATGTCCATGCC 268


RESULT 14
AB167637/c
LOCUS       AB167637                 398 bp    DNA     linear   VRL 01-OCT-2004
DEFINITION  Hepatitis B virus gene for polyprotein, partial cds, clone: NEP27.
ACCESSION   AB167637
VERSION     AB167637.1  GI:53148360
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1
  AUTHORS   Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
  TITLE     A Case-control Study for Differences among Hepatitis B Virus
            Infections of Genotypes A (Subtypes Aa and Ae) and D
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 398)
  AUTHORS   Tanaka,Y. and Mizokami,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-2004) Yasuhito Tanaka, Nagoya City University
            Graduate School of Medical Sciences, Department of Clinical
            Molecular Informative Medicine; 1 Kawasumi, Mizuho-cho, Mizuho-ku,
            Nagoya, Aichi 467-8601, Japan (E-mail:ytanaka@med.nagoya-cu.ac.jp,
            Tel:81-52-853-8292, Fax:81-52-842-0021)
FEATURES             Location/Qualifiers
     source          1. .398
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
                     /clone="NEP27"
     CDS             184. .>398
                     /note="precore-core region"
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="BAD52209.1"
                     /db_xref="GI:53148361"
                     /translation="MQLFHLCLIISCTCPTFQASKLCLGWLWGMDIDPYKEFGATVEL
                     LSFLPSDFFPSVRDLLDTASALYREAL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 398;
```

```
  Best Local Similarity   85.0%;  Pred. No. 18;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        287 TAAGGGTCGATGTCCATGCC 268


RESULT 15
AB163815/c
LOCUS       AB163815                406 bp    DNA     linear   VRL 29-JUN-2004
DEFINITION  Hepatitis B virus X, preC/C genes for polyproteins, isolate:
            SAF662.
ACCESSION   AB163815
VERSION     AB163815.1  GI:49387444
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1
  AUTHORS   Hasegawa,I., Tanaka,Y., Kramvis,A., Kato,T., Sugauchi,F.,
            Acharya,S.K., Orito,E., Ueda,R., Kew,M.C. and Mizokami,M.
  TITLE     Novel hepatitis B virus genotype a subtyping assay that
            distinguishes subtype aa from ae and its application in
            epidemiological studies
  JOURNAL   J. Virol. 78 (14), 7575-7581 (2004)
   PUBMED   15220432
REFERENCE   2  (bases 1 to 406)
  AUTHORS   Hasegawa,I., Tanaka,Y. and Mizokami,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-2004) Izumi Hasegawa, Nagoya City University
            Graduate School, Department of Internal Medicine and Molecular
            Science; 1 Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya, Aichi 467-8601,
            Japan (E-mail:izu-hase@med.nagoya-cu.ac.jp, Tel:81-52-853-8216,
            Fax:81-52-852-0849)
FEATURES             Location/Qualifiers
     source          1. .406
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="SAF662"
                     /db_xref="taxon:10407"
     gene            1. .208
                     /gene="X"
     CDS             <1. .208
                     /gene="X"
                     /codon_start=2
                     /product="polyprotein"
                     /protein_id="BAD24924.1"
                     /db_xref="GI:49387445"
                     /translation="QILPKVLHKRTLGLPAMSTTDLEAYFKDCVFKDWEELGEETRLM
                     IFVLGGCRHKLVCAPSSCNFFTSA"
     gene            184. .406
                     /gene="preC/C"
     CDS             184. .>406
                     /gene="preC/C"
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="BAD24925.1"
                     /db_xref="GI:49387446"
                     /translation="MQLFHLCLIISCTCPTVQASKLCLGWLWGMDIDPYKEFGATVEL
                     LSFLPQDFFPSVRDLLDTASALYREALESP"
ORIGIN

  Query Match            100.0%;  Score 20;  DB 14;  Length 406;
  Best Local Similarity   85.0%;  Pred. No. 18;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        287 TAAGGGTCGATGTCCATGCC 268


Search completed: March 17, 2005, 08:14:17
```

Job time : 683.733 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-63
Perfect score: 20
Sequence: 1 taagggucgauguccatgcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	2	AAT72572	Aat72572 Hepatitis
	2	20	100.0	20	2	AAT72573	Aat72573 Hepatitis
	3	20	100.0	30	2	AAT72619	Aat72619 Hepatitis
	4	20	100.0	30	2	AAT72620	Aat72620 Hepatitis
	5	20	100.0	30	2	AAT72618	Aat72618 Hepatitis
	6	20	100.0	30	2	AAT72621	Aat72621 Hepatitis
c	7	20	100.0	31	10	ADC64743	Adc64743 Hepatitis
c	8	20	100.0	639	6	AAD27422	Aad27422 Hepatitis
c	9	20	100.0	639	6	AAD31509	Aad31509 Hepatitis
c	10	20	100.0	663	3	AAA71734	Aaa71734 HBV fusio
c	11	20	100.0	669	12	ADO07220	Ado07220 Hepatitis
c	12	20	100.0	1334	2	AAV82691	Aav82691 Fulminant
c	13	20	100.0	1395	2	AAV82688	Aav82688 Fulminant
c	14	20	100.0	1400	2	AAV82687	Aav82687 Fulminant
c	15	20	100.0	1445	2	AAV82692	Aav82692 Fulminant
c	16	20	100.0	1445	2	AAV82685	Aav82685 Fulminant
c	17	20	100.0	1445	2	AAV82690	Aav82690 Fulminant
c	18	20	100.0	1445	2	AAV82684	Aav82684 Fulminant
c	19	20	100.0	1500	2	AAV82686	Aav82686 Fulminant
c	20	20	100.0	1500	2	AAV82689	Aav82689 Fulminant
c	21	20	100.0	2342	1	AAN93072	Aan93072 Sequence
c	22	20	100.0	3182	6	AAD31765	Aad31765 Hepatitis
c	23	20	100.0	3182	9	ACA62422	Aca62422 Hepatitis
c	24	20	100.0	3182	10	AAD60866	Aad60866 Hepatitis
c	25	20	100.0	5618	2	AAQ88310	Aaq88310 Plasmid p
c	26	20	100.0	7991	6	AAS16094	Aas16094 HBV viral
c	27	20	100.0	8007	6	AAS16092	Aas16092 HBV viral
c	28	20	100.0	8717	6	AAS16093	Aas16093 HBV viral
c	29	19	95.0	34	10	ADJ94544	Adj94544 SDMCore d
	30	19	95.0	34	10	ADJ94545	Adj94545 SDMCore d
c	31	19	95.0	39	13	ADR89273	Adr89273 Lab-on-ch
c	32	19	95.0	39	13	ADR89266	Adr89266 Lab-on-ch
c	33	18.8	94.0	30	2	AAQ45813	Aaq45813 HBV ampli
c	34	18.8	94.0	30	2	AAV07810	Aav07810 HBV.D46 a
c	35	18.8	94.0	30	2	AAV83039	Aav83039 Amplifier
	36	18.8	94.0	50	2	AAQ06723	Aaq06723 :HBV.LLA2
c	37	18.4	92.0	22	10	ADG46961	Adg46961 PCR prime
c	38	18.4	92.0	22	11	ADM83206	Adm83206 PCR prime
c	39	18.4	92.0	24	6	ABK44212	Abk44212 B cell ep
c	40	18.4	92.0	24	6	ABK67439	Abk67439 Primer #1
c	41	18.4	92.0	24	6	ABK67506	Abk67506 Hepatitis
c	42	18.4	92.0	24	10	ADE80023	Ade80023 Primer fo
c	43	18.4	92.0	24	10	ADE10976	Ade10976 Chimeric
c	44	18.4	92.0	24	10	ADG46965	Adg46965 PCR prime
c	45	18.4	92.0	24	11	ADM83210	Adm83210 PCR prime

ALIGNMENTS

RESULT 1
AAT72572
ID AAT72572 standard; DNA; 20 BP.
XX
AC AAT72572;
XX
DT 03-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV92b.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1. .20
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
XX
PN WO9639502-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP002432.
XX
PR 06-JUN-1995; 95US-00467397.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (HYBR-) HYBRIDON INC.
XX
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX
DR WPI; 1997-043124/04.
XX
PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
PS Claim 1; Page 12; 81pp; English.
XX
CC The present sequence represents a synthetic oligonucleotide HBV92b which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a

```
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 4 A; 5 C; 6 G; 5 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   85.0%;  Pred. No. 2.5;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db          1 TAAGGGTCGATGTCCATGCC 20


RESULT 2
AAT72573
ID   AAT72573 standard; DNA; 20 BP.
XX
AC   AAT72573;
XX
DT   03-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV92Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .20
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        11. .20
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   11
FT                   /*tag=  c
FT                   /mod_base= um
FT   modified_base   12
FT                   /*tag=  d
FT                   /mod_base= gm
FT   modified_base   13
FT                   /*tag=  e
FT                   /mod_base= um
FT   modified_base   14
FT                   /*tag=  f
FT                   /mod_base= cm
FT   modified_base   15
FT                   /*tag=  g
FT                   /mod_base= cm
FT   modified_base   16
FT                   /*tag=  h
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   17
FT                   /*tag=  i
FT                   /mod_base= um
FT   modified_base   18
FT                   /*tag=  j
FT                   /mod_base= gm
FT   modified_base   19
FT                   /*tag=  k
FT                   /mod_base= cm
FT   modified_base   20
FT                   /*tag=  l
FT                   /mod_base= cm
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV92Mb which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 4 A; 5 C; 6 G; 2 T; 3 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   90.0%;  Pred. No. 2.5;
  Matches   18;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:||||||||:|||
Db          1 TAAGGGTCGAUGUCCAUGCC 20


RESULT 3
AAT72619
ID   AAT72619 standard; DNA; 30 BP.
XX
AC   AAT72619;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-90Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        1. .10
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   1
FT                   /*tag=  c
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   2
FT                   /*tag=  d
FT                   /mod_base= gm
FT   modified_base   3
FT                   /*tag=  e
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   4
FT                   /*tag=  f
FT                   /mod_base= gm
FT   modified_base   5
FT                   /*tag=  g
```

```
FT                  /mod_base= OTHER
FT                  /note= "2'-O-methyladenosine"
FT   modified_base  6
FT                  /*tag=  h
FT                  /mod_base= um
FT   modified_base  7
FT                  /*tag=  i
FT                  /mod_base= gm
FT   modified_base  8
FT                  /*tag=  j
FT                  /mod_base= OTHER
FT                  /note= "2'-O-methyladenosine"
FT   modified_base  9
FT                  /*tag=  k
FT                  /mod_base= um
FT   modified_base  10
FT                  /*tag=  l
FT                  /mod_base= um
FT   misc_RNA       21. .30
FT                  /*tag=  m
FT                  /note= "2'-OMe RNA"
FT   modified_base  21
FT                  /*tag=  n
FT                  /mod_base= um
FT   modified_base  22
FT                  /*tag=  o
FT                  /mod_base= gm
FT   modified_base  23
FT                  /*tag=  p
FT                  /mod_base= um
FT   modified_base  24
FT                  /*tag=  q
FT                  /mod_base= cm
FT   modified_base  25
FT                  /*tag=  r
FT                  /mod_base= cm
FT   modified_base  26
FT                  /*tag=  s
FT                  /mod_base= OTHER
FT                  /note= "2'-O-methyladenosine"
FT   modified_base  27
FT                  /*tag=  t
FT                  /mod_base= um
FT   modified_base  28
FT                  /*tag=  u
FT                  /mod_base= gm
FT   modified_base  29
FT                  /*tag=  v
FT                  /mod_base= cm
FT   modified_base  30
FT                  /*tag=  w
FT                  /mod_base= cm
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-90Mb
CC   which contains a sequence which is complementary to at least two non-
CC   contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides complementary
CC   to a portion of the HBV RNA, may be used for inhibiting HBV replication
CC   in a cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 8 A; 5 C; 9 G; 2 T; 6 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   90.0%;  Pred. No. 2.6;
  Matches   18;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:||||||||:|||
Db         11 TAAGGGTCGAUGUCCAUGCC 30


RESULT 4
AAT72620
ID   AAT72620 standard; DNA; 30 BP.
XX
AC   AAT72620;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-91b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-91b which
CC   contains a sequence which is complementary to at least two non-contiguous
CC   regions of a hepatitis B virus (HBV) nucleic acid. The antisense
CC   oligonucleotide may be used to detect the presence of HBV in a sample.
CC   The antisense oligonucleotide, and oligonucleotides complementary to a
CC   portion of the HBV RNA, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 7 A; 6 C; 11 G; 6 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   85.0%;  Pred. No. 2.6;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db         11 TAAGGGTCGATGTCCATGCC 30


RESULT 5
AAT72618
ID   AAT72618 standard; DNA; 30 BP.
XX
AC   AAT72618;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-90b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-90b which
CC   contains a sequence which is complementary to at least two non-contiguous
CC   regions of a hepatitis B virus (HBV) nucleic acid. The antisense
CC   oligonucleotide may be used to detect the presence of HBV in a sample.
CC   The antisense oligonucleotide, and oligonucleotides complementary to a
CC   portion of the HBV RNA, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 30 BP; 8 A; 5 C; 9 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   85.0%;  Pred. No. 2.6;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db         11 TAAGGGTCGATGTCCATGCC 30


RESULT 6
AAT72621
ID   AAT72621 standard; DNA; 30 BP.
XX
AC   AAT72621;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV-91Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        21. .30
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   21
FT                   /*tag=  c
FT                   /mod_base= um
FT   modified_base   22
FT                   /*tag=  d
FT                   /mod_base= gm
FT   modified_base   23
FT                   /*tag=  e
FT                   /mod_base= um
FT   modified_base   24
FT                   /*tag=  f
FT                   /mod_base= cm
FT   modified_base   25
FT                   /*tag=  g
FT                   /mod_base= cm
FT   modified_base   26
FT                   /*tag=  h
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   27
FT                   /*tag=  i
FT                   /mod_base= um
FT   modified_base   28
FT                   /*tag=  j
FT                   /mod_base= gm
FT   modified_base   29
FT                   /*tag=  k
FT                   /mod_base= cm
FT   modified_base   30
FT                   /*tag=  l
FT                   /mod_base= cm
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 5; Page 15; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV-91Mb
CC   which contains a sequence which is complementary to at least two non-
CC   contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides complementary
CC   to a portion of the HBV RNA, may be used for inhibiting HBV replication
CC   in a cell or for the treatment of HBV infection
```

```
XX
SQ   Sequence 30 BP; 7 A; 6 C; 11 G; 3 T; 3 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   90.0%;  Pred. No. 2.6;
  Matches   18;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||||||||:|||
Db         11 TAAGGGTCGAUGUCCAUGCC 30


RESULT 7
ADC64743/c
ID   ADC64743 standard; RNA; 31 BP.
XX
AC   ADC64743;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Hepatitis B virus DNA polymerase related RNA oligonucleotide.
XX
KW   screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
XX
OS   Synthetic.
OS   Hepatitis B virus.
XX
PN   KR2002007891-A.
XX
PD   29-JAN-2002.
XX
PF   19-JUL-2000; 2000KR-00041420.
XX
PR   19-JUL-2000; 2000KR-00041420.
XX
PA   (MOGA-) MOGAM BIOTECHNOLOGY INST.
PA   (VIRO-) VIROGEN CO LTD.
XX
PI   Ji HJ,  Jung SI,  Kim YC,  Min MG,  Ryu WS,  Yoon GS;
XX
DR   WPI; 2003-309015/30.
XX
PT   Screening of antiviral agents by protein-priming activity of hepatitis B
PT   virus DNA polymerase.
XX
PS   Disclosure; Page 12; 13pp; Korean.
XX
CC   The present invention describes a method of screening for an antiviral
CC   agent by the protein-priming activity of hepatitis B virus (HBV) DNA
CC   polymerase. Also described is developing an antiviral agent with a high
CC   selectivity to HBV which can be used for high-throughput screening. The
CC   present sequence represents an RNA oligonucleotide which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 31 BP; 10 A; 6 C; 8 G; 0 T; 7 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 10;  Length 31;
  Best Local Similarity   85.0%;  Pred. No. 2.6;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db         22 TAAGGGTCGATGTCCATGCC 3


RESULT 8
AAD27422/c
ID   AAD27422 standard; DNA; 639 BP.
XX
AC   AAD27422;
XX
DT   18-APR-2002  (first entry)
XX
DE   Hepatitis B virus (HBV) core antigen (HBcAg) encoding DNA #1.
XX
KW   Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
KW   prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KW   hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW   tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KW   dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KW   food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
KW   antiprotozoal; ds.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag= a
FT                   /product= "HBcAg"
XX
PN   WO200198333-A2.
XX
PD   27-DEC-2001.
XX
PF   22-JUN-2001; 2001WO-GB002817.
XX
PR   22-JUN-2000; 2000GB-00015308.
PR   06-OCT-2000; 2000GB-00024544.
XX
PA   (CELL-) CELLTECH PHARM LTD.
XX
PI   Page M,  Li J,  Pumpens P;
XX
DR   WPI; 2002-098223/13.
DR   P-PSDB; AAE17018.
XX
PT   New proteins comprising a modified hepatitis B core antigen, useful as a
PT   vaccine in prophylactic or therapeutic vaccination of the human or animal
PT   body, particularly against hepatitis B virus infection.
XX
PS   Disclosure; Page 38-39; 40pp; English.
XX
CC   The invention relates to modified proteins comprising hepatitis B virus
CC   (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC   repeats has been deleted and the protein comprising the C-terminal
CC   cysteine of HBcAg. The deleted region may be replaced by an epitope from
CC   a protein other than HBcAg, in which case the HBcAg acts as a carrier to
CC   present the epitope to the immune system. This chimeric protein or its
CC   nucleic acid is useful as a vaccine or in a method of prophylactic or
CC   therapeutic vaccination of the human or animal body, particularly against
CC   HBV. The nucleic acid encoding the protein may be used in gene therapy or
CC   DNA vaccination protocols. The chimeric protein or its nucleic acid may
CC   also be used as the basis of a prophylactic vaccine against a range of
CC   diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
CC   influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
CC   immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
CC   tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
CC   diarrhoea, meningitis or gonorrhea. The present sequence is a DNA
CC   encoding Hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 6;  Length 639;
  Best Local Similarity   85.0%;  Pred. No. 3.3;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        104 TAAGGGTCGATGTCCATGCC 85


RESULT 9
AAD31509/c
ID   AAD31509 standard; DNA; 639 BP.
XX
AC   AAD31509;
```

```
XX
DT   18-JUN-2002  (first entry)
XX
DE   Hepatitis B virus core antigen (HBcAg) encoding DNA.
XX
KW   Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
KW   therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
KW   polio; herpes; rabies; AIDS; foot-and-mouth disease; ds.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .639
FT                   /*tag=  a
FT                   /product= "HBc protein"
FT   sig_peptide     1. .87
FT                   /*tag=  b
FT   mat_peptide     88. .636
FT                   /*tag=  c
FT                   /product= "Mature HBc protein"
XX
PN   WO200177158-A1.
XX
PD   18-OCT-2001.
XX
PF   09-APR-2001; 2001WO-GB001607.
XX
PR   07-APR-2000; 2000EP-00107118.
XX
PA   (MEDE-) MEDEVA EURO LTD.
XX
PI   Gehin A,  Gilbert R,  Stuart D,  Rowlands D;
XX
DR   WPI; 2002-239995/29.
DR   P-PSDB; AAE19793.
XX
PT   Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for the
PT   prophylactic or therapeutic treatment of humans or animals against e.g.
PT   HB virus, viral hepatitis, hepatitis C virus, influenza, or foot-and-
PT   mouth disease.
XX
PS   Disclosure; Page 23-24; 27pp; English.
XX
CC   The present invention relates to hepatitis B virus (HBV) core antigen
CC   (HBcAg) fusion proteins and polynucleotides encoding such proteins.
CC   Sequences of the invention are useful in methods of prophylactic or
CC   therapeutic vaccination or to manufacture medicaments for prophylactic or
CC   therapeutic vaccination of the human or animal body against HBV, e.g.
CC   against viral hepatitis. They are also useful as a prophylactic vaccine
CC   against e.g. hepatitis c virus, influenza, polio, herpes, rabies,
CC   acquired immune deficiency syndrome (AIDS) or foot-and-mouth disease. The
CC   present sequence is a DNA encoding hepatitis B virus core antigen (HBcAg)
XX
SQ   Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 6;  Length 639;
  Best Local Similarity   85.0%;  Pred. No. 3.3;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        104 TAAGGGTCGATGTCCATGCC 85


RESULT 10
AAA71734/c
ID   AAA71734 standard; cDNA; 663 BP.
XX
AC   AAA71734;
XX
DT   06-AUG-2003  (revised)
DT   08-JAN-2001  (first entry)
XX
DE   HBV fusion protein comprising LHB and RGD encoding cDNA.
XX
KW   Fusion protein; protein coat; virus-specific packaging signal; psi;
KW   virus protein; cell permeability; cell-specific binding site; LHB;
KW   large surface protein; core antigen; gene therapy; ss.
XX
OS   Hepatitis B virus.
OS   Synthetic.
XX
PN   WO200046376-A2.
XX
PD   10-AUG-2000.
XX
PF   04-FEB-2000; 2000WO-DE000363.
XX
PR   05-FEB-1999;   99DE-01004800.
XX
PA   (HILD/) HILDT E.
XX
PI   Hildt E,  Hofschneider P;
XX
DR   WPI; 2000-514959/46.
DR   P-PSDB; AAB10596.
XX
PT   Particle for cell-specific gene delivery, useful in gene therapy,
PT   comprises nucleic acid in protein coat that includes a fusion protein of
PT   viral protein, permeability peptide and cell-binding site.
XX
PS   Claim 16a; Fig 1; 34pp; German.
XX
CC   This invention describes a novel particle (A), comprising a protein coat
CC   with a fusion protein (FP), and, inside the coat, a nucleic acid (I)
CC   including the sequence for a virus-specific packaging signal (psi) and a
CC   structural gene. FP contains a virus protein (VP), a peptide (P) that
CC   mediates cell permeability and a heterologous cell-specific binding site
CC   (RGD). The invention also describes (1) producing (A) in which FP
CC   contains an LHBs (large surface protein of hepatitis B virus (HBV)) and
CC   RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),
CC   (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector
CC   containing the DNA of (d). The products of the invention are used in gene
CC   therapy of cells and tissues, in vivo or ex vivo. This sequence encodes a
CC   fusion protein which is described in the method of the invention.
CC   (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 663 BP; 154 A; 169 C; 152 G; 188 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 3;  Length 663;
  Best Local Similarity   85.0%;  Pred. No. 3.3;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        107 TAAGGGTCGATGTCCATGCC 88


RESULT 11
ADO07220/c
ID   ADO07220 standard; DNA; 669 BP.
XX
AC   ADO07220;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Hepatitis B virus core antigen DNA.
XX
KW   HBcAg; immunomodulator; vaccine; gene; ss.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             10. .669
FT                   /*tag=  a
FT                   /product= "HBcAg"
```

```
FT                   /partial
FT                   /note= "No start codon"
XX
PN   WO2004035007-A2.
XX
PD   29-APR-2004.
XX
PF   17-OCT-2003; 2003WO-US033178.
XX
PR   17-OCT-2002; 2002US-0419279P.
XX
PA   (ORAG-) ORAGEN CORP.
XX
PI   Michaels F;
XX
DR   WPI; 2004-348329/32.
DR   P-PSDB; ADO07221.
XX
PT   Modulating a systemic immune response to a peptide in a mammal comprises
PT   transmucosally administering a macromolecular aggregate of the peptide.
XX
PS   Disclosure; SEQ ID NO 1; 81pp; English.
XX
CC   The present sequence is the DNA sequence of the hepatitis B virus core
CC   antigen (HBcAg) gene from HBV serotype ayw. A peptide comprising a HBV
CC   protein can be used in claimed methods of the invention for modulating an
CC   immune response in a mammal. A method of inducing a systemic immune
CC   response to a peptide in a mammal comprises transmucosally administering
CC   to the mammal a macromolecular aggregate of the peptide. The
CC   macromolecular aggregate comprises at least 10 peptide subunits, may have
CC   a molecular weight of over 1,000 kDa, and is preferably at least 5 nm in
CC   diameter. It is resistant to digestive degradation, being stabilised in
CC   aggregate form by chemical treatment and/or by recombinant protein
CC   engineering of the peptide. The peptide preferably comprises a HBV
CC   protein selected from HBV surface protein, nucleocapsid protein or
CC   envelope protein. Transmucosal administration to a mammal of a
CC   macromolecular aggregate of a HBV surface protein engenders a systemic
CC   immune response in the mammal. A method of suppressing an immune response
CC   in a mammal involves transmucosally administering a monomolecular peptide
CC   that is resistant to digestive degradation and which may be stabilised by
CC   chemical treatment or protein engineering, and which may be derived from
CC   a HBV protein. A monomolecular peptide is useful for the induction of
CC   oral tolerance when induction of systemic immunity is undesirable, e.g.
CC   in cases of chronic infections.
XX
SQ   Sequence 669 BP; 155 A; 170 C; 148 G; 196 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 12;  Length 669;
  Best Local Similarity   85.0%;  Pred. No. 3.3;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        134 TAAGGGTCGATGTCCATGCC 115


RESULT 12
AAV82691/c
ID   AAV82691 standard; DNA; 1334 BP.
XX
AC   AAV82691;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Fulminant hepatitis B virus genotype D variant FHBV12 sequence.
XX
KW   Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW   HBV-related disease; ss.
XX
OS   Hepatitis B virus.
XX
PN   WO9845421-A2.
XX
PD   15-OCT-1998.
XX
PF   08-APR-1998;   98WO-EP002048.
XX
PR   09-APR-1997;   97GB-00007221.
XX
PA   (UNIU ) UNIV GLASGOW.
XX
PI   Carman B;
XX
DR   WPI; 1999-009329/01.
XX
PT   New hepatitis B virus nucleic acid with combination of specific mutations
PT   - useful for, e.g. detection of binding interactions between host or
PT   viral proteins and HBV nucleic.
XX
PS   Disclosure; Fig 5; 85pp; English.
XX
CC   The present sequence represents part of the genome of a fulminant
CC   Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC   The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC   a mutation (i.e. alteration from the normal nucleotide in any of the
CC   genotypes A to F) in at least two of the enhancer I region, the negative
CC   regulatory element region, the enhancer II/ core upstream regulatory
CC   sequence/ basal core promoter region, or a mutation which leads to an X-
CC   peptide amino acid change to Cys or Met. The HBV variants of the
CC   invention are used to detect binding interactions between host or viral
CC   proteins and HBV nucleic acid. Probes that hybridise to any of the
CC   specified mutated regions are used to detect HBV-related disease,
CC   especially fulminant infection, but also severe chronic infection or
CC   serologically unusual forms of disease. Combinations of the specified
CC   mutations are associated with fulminant infections, probably because they
CC   reduce the ability to bind inhibitory proteins in the host cell
XX
SQ   Sequence 1334 BP; 288 A; 363 C; 311 G; 372 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 1334;
  Best Local Similarity   85.0%;  Pred. No. 3.5;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        806 TAAGGGTCGATGTCCATGCC 787


RESULT 13
AAV82688/c
ID   AAV82688 standard; DNA; 1395 BP.
XX
AC   AAV82688;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Fulminant hepatitis B virus genotype D variant FHBV5 sequence.
XX
KW   Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW   HBV-related disease; ss.
XX
OS   Hepatitis B virus.
XX
PN   WO9845421-A2.
XX
PD   15-OCT-1998.
XX
PF   08-APR-1998;   98WO-EP002048.
XX
PR   09-APR-1997;   97GB-00007221.
XX
PA   (UNIU ) UNIV GLASGOW.
XX
PI   Carman B;
XX
DR   WPI; 1999-009329/01.
```

```
XX
PT   New hepatitis B virus nucleic acid with combination of specific mutations
PT   - useful for, e.g. detection of binding interactions between host or
PT   viral proteins and HBV nucleic.
XX
PS   Disclosure; Fig 5; 85pp; English.
XX
CC   The present sequence represents part of the genome of a fulminant
CC   Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC   The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC   a mutation (i.e. alteration from the normal nucleotide in any of the
CC   genotypes A to F) in at least two of the enhancer I region, the negative
CC   regulatory element region, the enhancer II/ core upstream regulatory
CC   sequence/ basal core promoter region, or a mutation which leads to an X-
CC   peptide amino acid change to Cys or Met. The HBV variants of the
CC   invention are used to detect binding interactions between host or viral
CC   proteins and HBV nucleic acid. Probes that hybridise to any of the
CC   specified mutated regions are used to detect HBV-related disease,
CC   especially fulminant infection, but also severe chronic infection or
CC   serologically unusual forms of disease. Combinations of the specified
CC   mutations are associated with fulminant infections, probably because they
CC   reduce the ability to bind inhibitory proteins in the host cell
XX
SQ   Sequence 1395 BP; 277 A; 387 C; 331 G; 398 T; 0 U; 2 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 1395;
  Best Local Similarity   85.0%;  Pred. No. 3.5;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        917 TAAGGGTCGATGTCCATGCC 898


RESULT 14
AAV82687/c
ID   AAV82687 standard; DNA; 1400 BP.
XX
AC   AAV82687;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Fulminant hepatitis B virus genotype D variant FHBV4 sequence.
XX
KW   Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW   HBV-related disease; ss.
XX
OS   Hepatitis B virus.
XX
PN   WO9845421-A2.
XX
PD   15-OCT-1998.
XX
PF   08-APR-1998;   98WO-EP002048.
XX
PR   09-APR-1997;   97GB-00007221.
XX
PA   (UNIU ) UNIV GLASGOW.
XX
PI   Carman B;
XX
DR   WPI; 1999-009329/01.
XX
PT   New hepatitis B virus nucleic acid with combination of specific mutations
PT   - useful for, e.g. detection of binding interactions between host or
PT   viral proteins and HBV nucleic.
XX
PS   Disclosure; Fig 5; 85pp; English.
XX
CC   The present sequence represents part of the genome of a fulminant
CC   Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC   The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC   a mutation (i.e. alteration from the normal nucleotide in any of the
CC   genotypes A to F) in at least two of the enhancer I region, the negative
CC   regulatory element region, the enhancer II/ core upstream regulatory
CC   sequence/ basal core promoter region, or a mutation which leads to an X-
CC   peptide amino acid change to Cys or Met. The HBV variants of the
CC   invention are used to detect binding interactions between host or viral
CC   proteins and HBV nucleic acid. Probes that hybridise to any of the
CC   specified mutated regions are used to detect HBV-related disease,
CC   especially fulminant infection, but also severe chronic infection or
CC   serologically unusual forms of disease. Combinations of the specified
CC   mutations are associated with fulminant infections, probably because they
CC   reduce the ability to bind inhibitory proteins in the host cell
XX
SQ   Sequence 1400 BP; 287 A; 388 C; 332 G; 393 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 1400;
  Best Local Similarity   85.0%;  Pred. No. 3.5;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||||||:|||:|:|||||||
Db        917 TAAGGGTCGATGTCCATGCC 898


RESULT 15
AAV82692/c
ID   AAV82692 standard; DNA; 1445 BP.
XX
AC   AAV82692;
XX
DT   16-FEB-1999  (first entry)
XX
DE   Fulminant hepatitis B virus genotype D variant FHBV13 sequence.
XX
KW   Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW   HBV-related disease; ss.
XX
OS   Hepatitis B virus.
XX
PN   WO9845421-A2.
XX
PD   15-OCT-1998.
XX
PF   08-APR-1998;   98WO-EP002048.
XX
PR   09-APR-1997;   97GB-00007221.
XX
PA   (UNIU ) UNIV GLASGOW.
XX
PI   Carman B;
XX
DR   WPI; 1999-009329/01.
XX
PT   New hepatitis B virus nucleic acid with combination of specific mutations
PT   - useful for, e.g. detection of binding interactions between host or
PT   viral proteins and HBV nucleic.
XX
PS   Disclosure; Fig 5; 85pp; English.
XX
CC   The present sequence represents part of the genome of a fulminant
CC   Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC   The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC   a mutation (i.e. alteration from the normal nucleotide in any of the
CC   genotypes A to F) in at least two of the enhancer I region, the negative
CC   regulatory element region, the enhancer II/ core upstream regulatory
CC   sequence/ basal core promoter region, or a mutation which leads to an X-
CC   peptide amino acid change to Cys or Met. The HBV variants of the
CC   invention are used to detect binding interactions between host or viral
CC   proteins and HBV nucleic acid. Probes that hybridise to any of the
CC   specified mutated regions are used to detect HBV-related disease,
CC   especially fulminant infection, but also severe chronic infection or
CC   serologically unusual forms of disease. Combinations of the specified
CC   mutations are associated with fulminant infections, probably because they
CC   reduce the ability to bind inhibitory proteins in the host cell
```

XX
SQ Sequence 1445 BP; 297 A; 406 C; 338 G; 404 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 1445;
Best Local Similarity 85.0%; Pred. No. 3.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAAGGGUCGAUGUCCATGCC 20
              |||||| ||| | |||||||
Db        917 TAAGGGTCGATGTCCATGCC 898
```

Search completed: March 17, 2005, 06:48:43
Job time : 171.333 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-63
Perfect score: 20
Sequence: 1 taagggucgauguccatgcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	19	95.0	686	4	BI954819	BI954819 HVSMEm001
	2	17.4	87.0	576	8	CC145068	CC145068 ZMMBBb000
	3	17.4	87.0	763	9	CL248902	CL248902 ZMMBBb059
c	4	17	85.0	260	7	CN589115	CN589115 TTE000135
	5	17	85.0	451	5	BP643309	BP643309 BP643309
c	6	17	85.0	729	4	BJ705222	BJ705222 BJ705222
	7	16.8	84.0	413	5	BP569691	BP569691 BP569691
	8	16.8	84.0	431	9	CL260820	CL260820 ZMMBBb061
	9	16.8	84.0	457	5	BP527907	BP527907 BP527907
	10	16.8	84.0	509	9	TA230D01P	AL481016 T. brucei
	11	16.8	84.0	733	1	AI635930	AI635930 tz82c11.x
c	12	16.8	84.0	773	8	BZ573222	BZ573222 msh2_3006
c	13	16.8	84.0	1574	9	AG476230	AG476230 Mus muscu
	14	16.4	82.0	272	2	BE922196	BE922196 EST425953
c	15	16.4	82.0	306	9	AG200800	AG200800 Pan trogl
	16	16.4	82.0	356	9	CC836806	CC836806 ZMMBBc048
c	17	16.4	82.0	417	7	H70178	H70178 yr94b08.r1
	18	16.4	82.0	499	5	BQ116265	BQ116265 EST601828
c	19	16.4	82.0	554	8	AQ433745	AQ433745 HS_5063_A
c	20	16.4	82.0	612	9	CL194160	CL194160 104_418_1
	21	16.4	82.0	622	5	BQ112668	BQ112668 EST598244
	22	16.4	82.0	633	9	CG461236	CG461236 PUFXP93TD
	23	16.4	82.0	678	9	CG761335	CG761335 ZMMBBb033
c	24	16.4	82.0	684	9	CL369255	CL369255 RPCI44_29
	25	16.4	82.0	714	9	AG128210	AG128210 Pan trogl
	26	16.4	82.0	772	8	BH257005	BH257005 CH230-283
c	27	16.4	82.0	848	7	CO099982	CO099982 GR_Ea25C
	28	16.4	82.0	851	4	BJ571152	BJ571152 BJ571152
	29	16.4	82.0	924	9	CG289400	CG289400 OGXBN55TV
	30	16.4	82.0	984	9	CG875596	CG875596 ZMMBBc028
	31	16.4	82.0	1010	9	CC622611	CC622611 OGULF56TV
c	32	16.4	82.0	1046	9	CL055525	CL055525 CH216-81E
	33	16.4	82.0	1669	3	HSM800819	AL110272 Homo sapi
	34	16.4	82.0	1669	3	HSM802167	AL137702 Homo sapi
	35	16	80.0	168	4	BI052244	BI052244 PM2-GN037
	36	16	80.0	168	4	BI052309	BI052309 PM2-GN037
	37	16	80.0	532	5	BX433931	BX433931 BX433931
c	38	16	80.0	666	9	CG106546	CG106546 PUIJJ59TB
c	39	16	80.0	698	8	BZ413874	BZ413874 if18d10.g
	40	16	80.0	740	9	CG235822	CG235822 OGWKA72TV
c	41	16	80.0	750	9	CG235810	CG235810 OGWKA72TH
c	42	16	80.0	810	9	CC716058	CC716058 OGTAI38TC
c	43	16	80.0	836	9	CG257272	CG257272 OG1CX72TH
	44	16	80.0	963	9	CG302315	CG302315 OG2BM75TV
	45	15.8	79.0	68	8	BH631190	BH631190 1007074G0

ALIGNMENTS

RESULT 1
BI954819/c
LOCUS BI954819 686 bp mRNA linear EST 19-OCT-2001
DEFINITION HVSMEm0019O17f Hordeum vulgare green seedling EST library HVcDNA0014 (Blumeria infected) Hordeum vulgare subsp. vulgare cDNA clone HVSMEm0019O17f, mRNA sequence.
ACCESSION BI954819
VERSION BI954819.1 GI:16300646
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 686)
AUTHORS Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Chin,A., Begum,D., Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Oates,R. and Main,D.
TITLE Development of a genetically and physically anchored EST resource for barley genomics: Blumeria infected Morex (compatible) seedling cDNA library
JOURNAL Unpublished (2001)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hq bases = 417
Seq primer: AATTAACCCTCACTAAAGGG
High quality sequence start: 16
High quality sequence stop: 531.
FEATURES Location/Qualifiers
source 1. .686
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSMEm0019O17f"
/tissue_type="green seedling leaf"
/lab_host="TJC121"
/clone_lib="Hordeum vulgare green seedling EST library HVcDNA0014 (Blumeria infected)"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Morex (mla) plants were greenhouse grown in the R

Wise lab at Iowa State University, Ames, IA; 7 day old green seedlings were infected with isolate 5874 of Blumeria graminis f. sp. hordei, and leaves were harvested 24, 48 and 72 hr post-inoculation and snap frozen (Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Chin). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

ORIGIN

```
  Query Match             95.0%;  Score 19;  DB 4;  Length 686;
  Best Local Similarity   84.2%;  Pred. No. 63;
  Matches   16;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AAGGGUCGAUGUCCATGCC 20
              |||||:|||:|:|||||||
Db        136 AAGGGTCGATGTCCATGCC 118
```

RESULT 2
CC145068

```
LOCUS       CC145068                576 bp    DNA     linear   GSS 24-JUN-2003
DEFINITION  ZMMBBb0002P18.r ZMMBBb Zea mays genomic clone ZMMBBb0002P18 3',
            genomic survey sequence.
ACCESSION   CC145068
VERSION     CC145068.1  GI:30090261
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 576)
  AUTHORS   Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
            and Wing,R.
  TITLE     Sequencing of the maize genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: T7
            BACKWARD: M13r
            Plate: 0002  row: P  column: 18
            Seq primer: M13r
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .576
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0002P18"
                     /lab_host="DH10B"
                     /clone_lib="ZMMBBb"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Zea mays L. ssp. mays"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 8;  Length 576;
  Best Local Similarity   84.2%;  Pred. No. 4.2e+02;
  Matches   16;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAGGGUCGAUGUCCATGCC 20
              |||||:|||:| |||||||
Db        139 AAGGGTCGATGCCCATGCC 157
```

RESULT 3
CL248902

```
LOCUS       CL248902                763 bp    DNA     linear   GSS 22-JAN-2004
DEFINITION  ZMMBBb0597N17r ZMMBBb (HindIII) Zea mays genomic clone
            ZMMBBb0597N17 3', genomic survey sequence.
ACCESSION   CL248902
VERSION     CL248902.1  GI:41105456
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 763)
  AUTHORS   Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
            Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
  TITLE     Sequencing of the maize genome at PGIR (2003c)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Bharti,A.K.
            Dr.Joachim Messing's lab
            The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
            University
            190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            Tel: 732 445 3801
            Fax: 732 445 5735
            Email: bharti@waksman.rutgers.edu
            Seq primer: SP6
            Class: BAC ends
            High quality sequence start: 405.
FEATURES             Location/Qualifiers
     source          1. .763
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0597N17"
                     /lab_host="E. coli DH10B"
                     /clone_lib="ZMMBBb (HindIII)"
                     /note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 763;
  Best Local Similarity   78.9%;  Pred. No. 4.4e+02;
  Matches   15;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAGGGUCGAUGUCCATGCC 20
              |||||:|||:|:| |||||
Db        339 AAGGGTCGATGTCTATGCC 357
```

RESULT 4
CN589115/c

```
LOCUS       CN589115                260 bp    mRNA    linear   EST 31-AUG-2004
DEFINITION  TTE00013587 Normalized large Tetrahymena thermophila cDNA, mRNA
            sequence.
ACCESSION   CN589115
```

```
VERSION     CN589115.1  GI:47040917
KEYWORDS    EST.
SOURCE      Tetrahymena thermophila
  ORGANISM  Tetrahymena thermophila
            Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
            Hymenostomatida; Tetrahymenina; Tetrahymenidae; Tetrahymena.
REFERENCE   1  (bases 1 to 260)
  AUTHORS   Garg,J., Pearlman,R.E. and Carlton,J.
  TITLE     PEPdbPub (http://amoebidia.bcm.umontreal.ca/public/pepdb/agrm.php)
            Tetrahymena thermophila (TIGR)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: PEPdb
            Departement de Biochimie, Universite de Montreal
            Email: pepdb-curator@bch.umontreal.ca
            Plate: 1398.
FEATURES             Location/Qualifiers
     source          1. .260
                     /organism="Tetrahymena thermophila"
                     /mol_type="mRNA"
                     /db_xref="taxon:5911"
                     /clone_lib="Normalized large"
ORIGIN

  Query Match            85.0%;  Score 17;  DB 7;  Length 260;
  Best Local Similarity  82.4%;  Pred. No. 6e+02;
  Matches   14;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GGGUCGAUGUCCATGCC 20
              |||:|||:|:|||||||
Db        237 GGGTCGATGTCCATGCC 221


RESULT 5
BP643309
LOCUS       BP643309                451 bp    mRNA    linear   EST 27-JUN-2004
DEFINITION  BP643309 RAFL19 Arabidopsis thaliana cDNA clone RAFL19-61-I08 3',
            mRNA sequence.
ACCESSION   BP643309
VERSION     BP643309.1  GI:49294779
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
            Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
            Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
  TITLE     Functional annotation of a full-length Arabidopsis cDNA collection
  JOURNAL   Science 296 (5565), 141-145 (2002)
  MEDLINE   21932900
   PUBMED   11910074
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            reversed clone; Please visit our web site
            (http://pfgweb.gsc.riken.go.jp/) for further details.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL19-61-I08"
                     /tissue_type="mixture of silique and flower"
                     /lab_host="DH10B"
                     /clone_lib="RAFL19"
                     /note="Site_1: BamHI; Site_2: SalI; Subtraction Library"
ORIGIN

  Query Match            85.0%;  Score 17;  DB 5;  Length 451;
  Best Local Similarity  82.4%;  Pred. No. 6.5e+02;
  Matches   14;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 AGGGUCGAUGUCCATGC 19
              ||||:|||:|:||||||
Db        431 AGGGTCGATGTCCATGC 447


RESULT 6
BJ705222/c
LOCUS       BJ705222                729 bp    mRNA    linear   EST 08-MAR-2004
DEFINITION  BJ705222 MF01FFA cDNA Oryzias latipes cDNA clone MF01FFA013a11 5',
            mRNA sequence.
ACCESSION   BJ705222
VERSION     BJ705222.1  GI:45246102
KEYWORDS    EST.
SOURCE      Oryzias latipes (Japanese medaka)
  ORGANISM  Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
  TITLE     Medaka EST Project in Takeda's lab
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Oryzias latipes"
                     /mol_type="mRNA"
                     /strain="Hd-rR"
                     /db_xref="taxon:8090"
                     /clone="MF01FFA013a11"
                     /sex="mixture of female and male"
                     /tissue_type="whole embryo"
                     /dev_stage="fry stage 40"
                     /clone_lib="MF01FFA cDNA"
ORIGIN

  Query Match            85.0%;  Score 17;  DB 4;  Length 729;
  Best Local Similarity  82.4%;  Pred. No. 7e+02;
  Matches   14;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AAGGGUCGAUGUCCATG 18
              |||||:|||:|:|||||
Db        182 AAGGGTCGATGTCCATG 166


RESULT 7
BP569691
LOCUS       BP569691                413 bp    mRNA    linear   EST 20-JUN-2004
DEFINITION  BP569691 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-68-M24 3',
            mRNA sequence.
ACCESSION   BP569691
VERSION     BP569691.1  GI:48985457
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
```

```
            Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
            Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
  TITLE     Functional annotation of a full-length Arabidopsis cDNA collection
  JOURNAL   Science 296 (5565), 141-145 (2002)
  MEDLINE   21932900
   PUBMED   11910074
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            reversed clone; Please visit our web site
            (http://pfgweb.gsc.riken.go.jp/) for further details.
FEATURES             Location/Qualifiers
     source          1. .413
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL14-68-M24"
                     /tissue_type="root"
                     /lab_host="DH10B"
                     /clone_lib="RAFL14"
                     /note="Site_1: BamHI; Site_2: SalI"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 5;  Length 413;
  Best Local Similarity   75.0%;  Pred. No. 8.2e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              |||||| ||:|:||||| |
Db        383 TAAGGGTTGATGTCCATGAC 402


RESULT 8
CL260820
LOCUS       CL260820                431 bp    DNA     linear   GSS 02-FEB-2004
DEFINITION  ZMMBBb0619B24r ZMMBBb (HindIII) Zea mays genomic clone
            ZMMBBb0619B24 3', genomic survey sequence.
ACCESSION   CL260820
VERSION     CL260820.1  GI:41624959
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
            Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
  TITLE     Sequencing of the maize genome at PGIR (2003c)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Bharti,A.K.
            Dr.Joachim Messing's lab
            The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
            University
            190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            Tel: 732 445 3801
            Fax: 732 445 5735
            Email: bharti@waksman.rutgers.edu
            Seq primer: SP6
            Class: BAC ends
            High quality sequence start: 116.
FEATURES             Location/Qualifiers
     source          1. .431
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0619B24"
                     /lab_host="E. coli DH10B"
                     /clone_lib="ZMMBBb (HindIII)"
                     /note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 431;
  Best Local Similarity   80.0%;  Pred. No. 8.3e+02;
  Matches   16;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              |||||:|||:| ||| |||
Db        328 TAAGGGTCGATGCCCAGGCC 347


RESULT 9
BP527907
LOCUS       BP527907                457 bp    mRNA    linear   EST 28-SEP-2004
DEFINITION  BP527907 MAT001 Nicotiana tabacum cDNA clone BY12728, mRNA
            sequence.
ACCESSION   BP527907
VERSION     BP527907.1  GI:52831634
KEYWORDS    EST.
SOURCE      Nicotiana tabacum (common tobacco)
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Matsuoka,K., Tashiro,G., Horiguchi,T., Demura,T. and Fukuda,H.
  TITLE     Profiling growth-phase dependent gene expression of tobacco BY-2
            cells by comprehensive microarray analysis
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ken Matsuoka
            Morphogenesis Research Group
            RIKEN Plant Science Center
            1-7-2 Suehirocho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9575
            Fax: 81-45-503-9573
            Email: by2@psc.riken.go.jp, URL:http://mrg.psc.riken.go.jp/strc/
            The cDNA library was constructed from mRNA isolated from lag (9 h),
            log (72 h) and stationary (7 days) old BY-2 cells.
            Seq primer: M13 forward.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /cultivar="Bright Yellow No.2"
                     /db_xref="taxon:4097"
                     /clone="BY12728"
                     /cell_line="BY-2"
                     /clone_lib="MAT001"
                     /note="Vector: pGEM-T easy; primer: M13 forward; mRNA
                     obtained from lag, log and stationary phase cells"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 5;  Length 457;
  Best Local Similarity   75.0%;  Pred. No. 8.3e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||| | :|||:|:|||||||
Db        297 TAAAGTTCGATGTCCATGCC 316


RESULT 10
TA230D01P
LOCUS       TA230D01P               509 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 230d01, forward sequence,
            genomic survey sequence.
ACCESSION   AL481016
VERSION     AL481016.1  GI:11846785
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
```

```
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 509)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb). The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects. In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES              Location/Qualifiers
     source           1. .509
                      /organism="Trypanosoma brucei"
                      /mol_type="genomic DNA"
                      /strain="TREU927"
                      /db_xref="taxon:5691"
                      /clone="230d01"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 509;
  Best Local Similarity   75.0%;  Pred. No. 8.5e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              |||||| ||:|:||| |||
Db        355 TAAGGGTTGATGTCCAGGCC 374


RESULT 11
AI635930
LOCUS       AI635930                733 bp    mRNA    linear   EST 16-DEC-1999
DEFINITION  tz82c11.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2295092 3'
            similar to gb:J03490 DIHYDROLIPOAMIDE DEHYDROGENASE PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   AI635930
VERSION     AI635930.1  GI:4687260
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 733)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11548-013
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1107   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 401.
FEATURES              Location/Qualifiers
     source           1. .733
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2295092"
                      /tissue_type="adenocarcinoma"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Pan1"
                      /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.72 kb. Life Technologies catalog #:
                      11548-013"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 1;  Length 733;
  Best Local Similarity   75.0%;  Pred. No. 8.9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              |||||| ||:|:||| |||| |
Db        557 TAAGGGTCGATGTGCATGAC 576


RESULT 12
BZ573222/c
LOCUS       BZ573222                773 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  msh2_3006.y2 msh Pseudomonas aeruginosa genomic clone msh2_3006,
            genomic survey sequence.
ACCESSION   BZ573222
VERSION     BZ573222.1  GI:27208283
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES              Location/Qualifiers
     source           1. .773
                      /organism="Pseudomonas aeruginosa"
                      /mol_type="genomic DNA"
                      /strain="MSH"
                      /db_xref="taxon:287"
                      /clone="msh2_3006"
                      /clone_lib="msh"
                      /note="Environmental isolate. Whole genomic shotgun
                      library."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 8;  Length 773;
  Best Local Similarity   75.0%;  Pred. No. 9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              ||| ||:|||:|:||| |||
Db        294 TAATGGTCGATGTCCAGGCC 275


RESULT 13
AG476230/c
LOCUS       AG476230                1574 bp    DNA     linear   GSS 04-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-369C21.TJ, genomic survey
            sequence.
```

```
ACCESSION   AG476230
VERSION     AG476230.1  GI:48183460
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1574)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1574
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-369C21.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 1574;
  Best Local Similarity   75.0%;  Pred. No. 1e+03;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TAAGGGUCGAUGUCCATGCC 20
              || |||:|||:|:|||| ||
Db        300 TATGGGTCGATGTCCATCCC 281


RESULT 14
BE922196
LOCUS       BE922196                272 bp    mRNA    linear   EST 07-MAR-2003
DEFINITION  EST425953 potato leaves and petioles Solanum tuberosum cDNA clone
            cSTB18G4 5' sequence, mRNA sequence.
ACCESSION   BE922196
VERSION     BE922196.1  GI:10448260
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 272)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
            Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R.,
            Ronning,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
  TITLE     Generation of ESTs from potato leaves and petioles
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/.
FEATURES             Location/Qualifiers
     source          1. .272
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="cSTB18G4"
                     /tissue_type="leaflets and petioles"
                     /dev_stage="8 weeks old plants"
                     /lab_host="SOLR"
                     /clone_lib="potato leaves and petioles"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; Tissue was supplied by Dr. Fry (Cornell University).
                     Leaflets and petioles were isolated from 8 week old
                     greenhouse grown plants. The plants were watered and
                     fertilized freely. The tissue was immediately frozen in
                     liquid nitrogen."
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 2;  Length 272;
  Best Local Similarity   77.8%;  Pred. No. 1.2e+03;
  Matches   14;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 AGGGUCGAUGUCCATGCC 20
              ||| :|||:|:|||||||
Db         29 AGGTTCGATGTCCATGCC 46


RESULT 15
AG200800/c
LOCUS       AG200800                306 bp    DNA     linear   GSS 06-MAR-2004
DEFINITION  Pan troglodytes DNA, clone: RP43-082P16.T7, genomic survey
            sequence.
ACCESSION   AG200800
VERSION     AG200800.1  GI:45232975
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     BAC end sequences of Library RP-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 306)
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of
            Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC);
            52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea
            (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/,
            Tel:82-42-866-7181, Fax:82-42-860-4409)
COMMENT     Clones are derived from the chimpanzee BAC library RP-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
            PRIMERS
                Sequencing: T7
            LIBRARY
                Vector    : pBACe3.6
                R.Site 1  : EcoRI
                R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .306
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
```

/db_xref="taxon:9598"
/clone="RP43-082P16.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RP-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 82.0%; Score 16.4; DB 9; Length 306;
Best Local Similarity 77.8%; Pred. No. 1.3e+03;
Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TAAGGGUCGAUGUCCATG 18
              |||||| ||:|:|||||
Db         35 TAAGGGTAGATGTCCATG 18
```

Search completed: March 17, 2005, 11:07:49
Job time : 1386.27 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-64
Perfect score: 20
Sequence: 1 ttataagggtcgauguccau 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	6	AR027818	AR027818 Sequence
2	20	100.0	21	6	AX137490	AX137490 Sequence
3	20	100.0	21	6	BD011606	BD011606 Biologica
4	20	100.0	27	6	AR027819	AR027819 Sequence
c 5	20	100.0	28	6	AR103926	AR103926 Sequence
c 6	20	100.0	35	6	BD236992	BD236992 DNA vacci
c 7	20	100.0	81	6	I92348	I92348 Sequence 9
c 8	20	100.0	253	14	AY329529	AY329529 Hepatitis
c 9	20	100.0	253	14	AY329561	AY329561 Hepatitis
c 10	20	100.0	253	14	AY329562	AY329562 Hepatitis
c 11	20	100.0	253	14	AY329568	AY329568 Hepatitis
c 12	20	100.0	253	14	AY329573	AY329573 Hepatitis
c 13	20	100.0	253	14	AY329575	AY329575 Hepatitis
c 14	20	100.0	253	14	AY329581	AY329581 Hepatitis
c 15	20	100.0	294	14	AF390000	AF390000 Hepatitis
c 16	20	100.0	333	14	HPBHBED	L12359 Hepatitis B
c 17	20	100.0	398	14	AB167603	AB167603 Hepatitis
c 18	20	100.0	398	14	AB167637	AB167637 Hepatitis
c 19	20	100.0	406	14	AB163815	AB163815 Hepatitis
c 20	20	100.0	406	14	AB163817	AB163817 Hepatitis
c 21	20	100.0	439	14	AY254503	AY254503 Hepatitis
c 22	20	100.0	456	14	AY509204	AY509204 Hepatitis
c 23	20	100.0	488	14	AY274419	AY274419 Hepatitis
c 24	20	100.0	488	14	AY274420	AY274420 Hepatitis
c 25	20	100.0	488	14	AY274422	AY274422 Hepatitis
c 26	20	100.0	488	14	AY274427	AY274427 Hepatitis
c 27	20	100.0	488	14	AY274428	AY274428 Hepatitis
c 28	20	100.0	488	14	AY274429	AY274429 Hepatitis
c 29	20	100.0	488	14	AY274430	AY274430 Hepatitis
c 30	20	100.0	488	14	AY274431	AY274431 Hepatitis
c 31	20	100.0	488	14	AY274432	AY274432 Hepatitis
c 32	20	100.0	488	14	AY274433	AY274433 Hepatitis
c 33	20	100.0	488	14	AY274434	AY274434 Hepatitis
c 34	20	100.0	488	14	AY274436	AY274436 Hepatitis
c 35	20	100.0	548	14	AY382500	AY382500 Hepatitis
c 36	20	100.0	548	14	AY382501	AY382501 Hepatitis
c 37	20	100.0	548	14	AY382502	AY382502 Hepatitis
c 38	20	100.0	548	14	AY382521	AY382521 Hepatitis
c 39	20	100.0	548	14	AY382522	AY382522 Hepatitis
c 40	20	100.0	548	14	AY382523	AY382523 Hepatitis
c 41	20	100.0	548	14	AY382524	AY382524 Hepatitis
c 42	20	100.0	548	14	AY382525	AY382525 Hepatitis
c 43	20	100.0	548	14	AY382526	AY382526 Hepatitis
c 44	20	100.0	548	14	AY382527	AY382527 Hepatitis
c 45	20	100.0	552	6	BD236991	BD236991 DNA vacci

ALIGNMENTS

RESULT 1
AR027818
LOCUS AR027818 20 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 16 from patent US 5856459.
ACCESSION AR027818
VERSION AR027818.1 GI:5938638
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and Mills,J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 16 05-JAN-1999;
FEATURES Location/Qualifiers
source 1. .20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 85.0%; Pred. No. 2;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATAAGGGTCGAUGUCCAU 20
||||||||||||:|:|||:
Db 1 TTATAAGGGTCGATGTCCAT 20

RESULT 2
AX137490
LOCUS AX137490 21 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 3 from Patent EP1072271.
ACCESSION AX137490
VERSION AX137490.1 GI:14273684
KEYWORDS .
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE 1
AUTHORS Rabbani,E., Ilan,Y., Roy-Chowdhury,J. and Engelhardt,D.L.

```
  TITLE     Selective immune down regulation (sidr) mediated transplantation
            processes
  JOURNAL   Patent: EP 1072271-A 3 31-JAN-2001;
            ENZO THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Hepatitis B virus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10407"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 21;
  Best Local Similarity   85.0%;  Pred. No. 2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db          2 TTATAAGGGTCGATGTCCAT 21


RESULT 3
BD011606
LOCUS       BD011606                21 bp    DNA     linear   PAT 09-JAN-2004
DEFINITION  Biological models showing secondary disease sign and useful in
            developing remedies, diagnostic products and therapeutic or
            diagnostic procedure, method with the use of the same and cells,
            tissues and organs derived therefrom.
ACCESSION   BD011606
VERSION     BD011606.1  GI:18639979
KEYWORDS    JP 2001078621-A/3.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Brown,J.J., Rabbani,I., Donegan,J.J. and Chaudhurry,J.R.
  TITLE     Biological models showing secondary disease sign and useful in
            developing remedies, diagnostic products and therapeutic or
            diagnostic procedure, method with the use of the same and cells,
            tissues and organs derived therefrom
  JOURNAL   Patent: JP 2001078621-A 3 27-MAR-2001;
            ENZO THERAPEUTICS INC
COMMENT     OS   Hepatitis virus (hepatitis B virus)
            PN   JP 2001078621-A/3
            PD   27-MAR-2001
            PF   14-JUL-2000 JP 2000215182
            PR   16-JUL-1999 US    09/356293
            PI   JENNIFER JUNE BROWN,IRAZAR RABBANI,JAMES J DONEGAN, PI
            JAYANTA ROY CHAUDHURRY
            PC   A01K67/027,A61K45/00,C12N5/00,C12N15/09,C12Q1/68,G01N33/15, PC
              G01N33/50//
            PC   C12N1/20,C12N7/00,C12N5/00,C12N15/00
            CC
            FH   Key             Location/Qualifiers
            FT   source          1. .21
            FT   /organism='Hepatitis virus (hepatitis B virus)'.
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 21;
  Best Local Similarity   85.0%;  Pred. No. 2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db          2 TTATAAGGGTCGATGTCCAT 21


RESULT 4
AR027819
LOCUS       AR027819                27 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 17 from patent US 5856459.
ACCESSION   AR027819
VERSION     AR027819.1  GI:5938639
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and
            Mills,J.S.
  TITLE     Oligonucleotides specific for hepatitis B virus
  JOURNAL   Patent: US 5856459-A 17 05-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 27;
  Best Local Similarity   85.0%;  Pred. No. 2.1;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db          8 TTATAAGGGTCGATGTCCAT 27


RESULT 5
AR103926/c
LOCUS       AR103926                28 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 1 from patent US 6087556.
ACCESSION   AR103926
VERSION     AR103926.1  GI:12815514
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Feitelson,M. and Siracusa,L.
  TITLE     Transgenic animals capable of replicating hepatitis viruses and
            mimicking chronic hepatitis infection in humans
  JOURNAL   Patent: US 6087556-A 1 11-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 28;
  Best Local Similarity   85.0%;  Pred. No. 2.1;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db         20 TTATAAGGGTCGATGTCCAT 1


RESULT 6
BD236992/c
LOCUS       BD236992                35 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  DNA vaccination to cholesterol ester transfer protein in the
            treatment of atherosclerosis.
ACCESSION   BD236992
VERSION     BD236992.1  GI:33046762
KEYWORDS    JP 2002516656-A/17.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Needleman,P. and Glenn,K.
  TITLE     DNA vaccination to cholesterol ester transfer protein in the
            treatment of atherosclerosis
```

```
  JOURNAL   Patent: JP 2002516656-A 17 11-JUN-2002;
            MONSANTO CO
COMMENT     OS   Unidentified
            PN   JP 2002516656-A/17
            PD   11-JUN-2002
            PF   17-SEP-1998 JP 2000512947
            PR   19-SEP-1997 US    08/934367
            PI   PHILIP NEEDLEMAN,KEVIN GLENN
            PC   C12N15/09,A61K48/00,C12N15/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   DNA vaccination to cholesterol ester transfer protein in the
            CC           treatment of
            CC   atherosclerosis
            FH   Key             Location/Qualifiers
            FT   source          1. .35
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 35;
  Best Local Similarity   85.0%;  Pred. No. 2.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||||:|:|||:
Db         26 TTATAAGGGTCGATGTCCAT 7


RESULT 7
I92348/c
LOCUS       I92348                   81 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 9 from patent US 5728518.
ACCESSION   I92348
VERSION     I92348.1  GI:3936818
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Carmichael,E.
  TITLE     Antiviral poly-and oligonucleotides
  JOURNAL   Patent: US 5728518-A 9 17-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .81
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 81;
  Best Local Similarity   85.0%;  Pred. No. 2.5;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||||:|:|||:
Db         74 TTATAAGGGTCGATGTCCAT 55


RESULT 8
AY329529/c
LOCUS       AY329529                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate A611252E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329529
VERSION     AY329529.1  GI:37625315
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="A611252E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95861.1"
                     /db_xref="GI:37625316"
                     /translation="STTDLEAYFKDCVFKDWEELGEETRLMIFVLGGCRHKLVCAPAS
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95862.1"
                     /db_xref="GI:37625317"
                     /translation="MQLFHLCLVISCTCPTFQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 2.9;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||||:|:|||:
Db        240 TTATAAGGGTCGATGTCCAT 221


RESULT 9
AY329561/c
LOCUS       AY329561                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D272811E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329561
VERSION     AY329561.1  GI:37625410
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
```

```
                     /isolate="D272811E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95924.1"
                     /db_xref="GI:37625411"
                     /translation="STTDLEAYFKDCLFKDWEELGEETRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95925.1"
                     /db_xref="GI:37625412"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 2.9;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||||||||||||:|:|||:
Db        240 TTATAAGGGTCGATGTCCAT 221


RESULT 10
AY329562/c
LOCUS       AY329562                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D273984E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329562
VERSION     AY329562.1  GI:37625413
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D273984E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95926.1"
                     /db_xref="GI:37625414"
                     /translation="STTDLEAYFKDCLFKDWEELGEEIRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95927.1"
                     /db_xref="GI:37625415"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 2.9;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||||||||||||:|:|||:
Db        240 TTATAAGGGTCGATGTCCAT 221


RESULT 11
AY329568/c
LOCUS       AY329568                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D296668E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329568
VERSION     AY329568.1  GI:37625431
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D296668E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95938.1"
                     /db_xref="GI:37625432"
                     /translation="STTDLEAYFKDCLFKDWEELGEELRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95939.1"
                     /db_xref="GI:37625433"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 2.9;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||||||||||||:|:|||:
Db        240 TTATAAGGGTCGATGTCCAT 221


RESULT 12
AY329573/c
LOCUS       AY329573                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D604917E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329573
VERSION     AY329573.1  GI:37625446
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
```

```
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D604917E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95948.1"
                     /db_xref="GI:37625447"
                     /translation="STTDLEAYFKDCLFKDWEELGEEIRLKVFVLGGCRHKLVCAPAS
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95949.1"
                     /db_xref="GI:37625448"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 2.9;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||||:|:|||:
Db        240 TTATAAGGGTCGATGTCCAT 221


RESULT 13
AY329575/c
LOCUS       AY329575                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D611058E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329575
VERSION     AY329575.1  GI:37625452
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D611058E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95952.1"
                     /db_xref="GI:37625453"
                     /translation="STTDLEAYFKDCLFKDWEELGEETRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95953.1"
                     /db_xref="GI:37625454"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 2.9;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||||:|:|||:
Db        240 TTATAAGGGTCGATGTCCAT 221


RESULT 14
AY329581/c
LOCUS       AY329581                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D639472E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329581
VERSION     AY329581.1  GI:37625470
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D639472E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95964.1"
                     /db_xref="GI:37625471"
                     /translation="STTDLEAYFKDCLFKDWEELGEELRLLIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95965.1"
                     /db_xref="GI:37625472"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN
```

```
  Query Match            100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   85.0%;  Pred. No. 2.9;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db        240 TTATAAGGGTCGATGTCCAT 221


RESULT 15
AF390000/c
LOCUS       AF390000                294 bp    DNA     linear   VRL 06-MAR-2002
DEFINITION  Hepatitis B virus isolate D3 X protein gene, partial cds; and
            nonfunctional precore/core protein gene, partial sequence.
ACCESSION   AF390000
VERSION     AF390000.1  GI:16266099
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   Castro,L.D., Niel,C. and Gomes,S.A.
  TITLE     Low frequency of mutations in the core promoter and precore regions
            of hepatitis B virus in anti-HBe positive Brazilian carriers
  JOURNAL   BMC Microbiol. 1 (1), 10 (2001)
   PUBMED   11472634
REFERENCE   2  (bases 1 to 294)
  AUTHORS   De Castro,L., Niel,C. and Gomes,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-2001) Virology, FIOCRUZ, Av. Brasil 4365, Rio de
            Janeiro, RJ 21045-900, Brazil
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D3"
                     /db_xref="taxon:10407"
     CDS             <1. .119
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAL16752.1"
                     /db_xref="GI:16266100"
                     /translation="FKDWEELGDDSRLMIYVLGGCRHKLVCAPAPCNFFTSA"
     misc_feature    95. .>294
                     /note="nonfunctional precore/core protein due to mutation"
ORIGIN

  Query Match            100.0%;  Score 20;  DB 14;  Length 294;
  Best Local Similarity   85.0%;  Pred. No. 3;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db        201 TTATAAGGGTCGATGTCCAT 182


Search completed: March 17, 2005, 08:14:18
Job time : 684.733 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-64
Perfect score: 20
Sequence: 1 ttataagggtcgauguccau 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	2	AAT72575	Aat72575 Hepatitis
	2	20	100.0	20	2	AAT72576	Aat72576 Hepatitis
	3	20	100.0	21	2	AAQ38441	Aaq38441 Antisense
	4	20	100.0	21	3	AAZ98697	Aaz98697 Human hep
	5	20	100.0	21	5	AAF58672	Aaf58672 Hepatitis
	6	20	100.0	21	10	ADB68574	Adb68574 NG2 A-L-P
	7	20	100.0	21	10	ADB68572	Adb68572 A-L-P con
	8	20	100.0	27	2	AAT72577	Aat72577 Hepatitis
c	9	20	100.0	28	2	AAQ80499	Aaq80499 Primer to
c	10	20	100.0	28	2	AAV45201	Aav45201 Primer MF
c	11	20	100.0	28	3	AAA62585	Aaa62585 Transgeni
c	12	20	100.0	28	12	ADN36070	Adn36070 Probe #15
c	13	20	100.0	28	12	ADN36068	Adn36068 Probe #14
c	14	20	100.0	28	12	ADN36071	Adn36071 Probe #15
c	15	20	100.0	28	12	ADN36073	Adn36073 Probe #15
c	16	20	100.0	28	12	ADN36066	Adn36066 Probe #14
c	17	20	100.0	28	12	ADN36067	Adn36067 Probe #14
c	18	20	100.0	28	12	ADN36072	Adn36072 Probe #15
c	19	20	100.0	28	12	ADN36065	Adn36065 Probe #14
c	20	20	100.0	28	12	ADN36069	Adn36069 Probe #15
c	21	20	100.0	31	6	ABA96791	Aba96791 Hepatitis
c	22	20	100.0	31	10	ADC64743	Adc64743 Hepatitis
c	23	20	100.0	35	2	AAX36590	Aax36590 PCR prime
c	24	20	100.0	35	8	ABX95880	Abx95880 PCR prime
c	25	20	100.0	35	10	ACD07807	Acd07807 Hepatitis
c	26	20	100.0	36	10	ADJ94539	Adj94539 HBV genom
c	27	20	100.0	39	13	ADR89273	Adr89273 Lab-on-ch
c	28	20	100.0	39	13	ADR89266	Adr89266 Lab-on-ch
c	29	20	100.0	53	12	ADN36055	Adn36055 Probe #13
	30	20	100.0	504	11	ADM41005	Adm41005 HBc relat
c	31	20	100.0	504	11	ADM41004	Adm41004 HBc relat
c	32	20	100.0	513	6	ABK67524	Abk67524 DNA encod
c	33	20	100.0	513	6	ABK67525	Abk67525 DNA encod
c	34	20	100.0	516	6	ABK67527	Abk67527 DNA encod
c	35	20	100.0	519	6	ABK67526	Abk67526 DNA encod
c	36	20	100.0	534	11	ADM40998	Adm40998 HBc relat
	37	20	100.0	534	11	ADM40999	Adm40999 HBc relat
c	38	20	100.0	534	11	ADM41007	Adm41007 HBc relat
	39	20	100.0	534	11	ADM41008	Adm41008 HBc relat
c	40	20	100.0	540	11	ADM41010	Adm41010 HBc relat
	41	20	100.0	540	11	ADM41011	Adm41011 HBc relat
c	42	20	100.0	549	6	ABK44278	Abk44278 DNA encod
c	43	20	100.0	549	6	ABK67533	Abk67533 Immunogen
c	44	20	100.0	549	10	ADE10968	Ade10968 Human Hep
c	45	20	100.0	549	10	ADG47010	Adg47010 Hepatitis

ALIGNMENTS

RESULT 1
AAT72575
ID AAT72575 standard; DNA; 20 BP.
XX
AC AAT72575;
XX
DT 04-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV101b.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1. .20
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
XX
PN WO9639502-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP002432.
XX
PR 06-JUN-1995; 95US-00467397.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (HYBR-) HYBRIDON INC.
XX
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX
DR WPI; 1997-043124/04.
XX
PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
PS Claim 1; Page 12; 81pp; English.
XX
CC The present sequence represents a synthetic oligonucleotide HBV101b which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a

```
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 5 A; 3 C; 5 G; 7 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   85.0%;  Pred. No. 1.1;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db          1 TTATAAGGGTCGATGTCCAT 20


RESULT 2
AAT72576
ID   AAT72576 standard; DNA; 20 BP.
XX
AC   AAT72576;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV101Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .20
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        14. .20
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   14
FT                   /*tag=  c
FT                   /mod_base= um
FT   modified_base   15
FT                   /*tag=  d
FT                   /mod_base= gm
FT   modified_base   16
FT                   /*tag=  e
FT                   /mod_base= um
FT   modified_base   17
FT                   /*tag=  f
FT                   /mod_base= cm
FT   modified_base   18
FT                   /*tag=  g
FT                   /mod_base= cm
FT   modified_base   19
FT                   /*tag=  h
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
FT   modified_base   20
FT                   /*tag=  i
FT                   /mod_base= um
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV101Mb
CC   which is complementary to a portion of the hepatitis B virus (HBV) RNA.
CC   The antisense oligonucleotide may be used to detect the presence of HBV
CC   in a sample. The antisense oligonucleotide, and oligonucleotides
CC   containing a sequence which is complementary to at least two non-
CC   contiguous regions of an HBV nucleic acid, may be used for inhibiting HBV
CC   replication in a cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 5 A; 3 C; 5 G; 4 T; 3 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||||||||||
Db          1 TTATAAGGGTCGAUGUCCAU 20


RESULT 3
AAQ38441
ID   AAQ38441 standard; DNA; 21 BP.
XX
AC   AAQ38441;
XX
DT   25-MAR-2003  (revised)
DT   08-JUL-1993  (first entry)
XX
DE   Antisense oligomer hybridising to HBV poly A site.
XX
KW   Hepatitis B virus; polyadenylation; targetting; cell-specific; complex;
KW   oncogenes; pathogen; parasite; ss.
XX
OS   Synthetic.
XX
PN   WO9304701-A1.
XX
PD   18-MAR-1993.
XX
PF   04-SEP-1992;   92WO-US007339.
XX
PR   05-SEP-1991;   91US-00755083.
PR   04-NOV-1991;   91US-00788119.
PR   03-APR-1992;   92US-00864003.
XX
PA   (UYCO-) UNIV CONNECTICUT.
XX
PI   Wu GY;
XX
DR   WPI; 1993-100661/12.
XX
PT   Soluble mol. complex for targetting delivery of poly- or oligo-
PT   nucleotide(s) to cells - includes carrier comprising cell-specific
PT   binding agent and poly- or oligo-nucleotide-binding agent.
XX
PS   Claim 38; Page 12; 44pp; English.
XX
CC   The sequence corresponds to nucleotides 1903-1923 of the Hepatitis B
CC   viral genome (the polyadenylation site). The oligonucleotide was used to
CC   specifically hybridise with HBV in a targettable, soluble DNA soln., to
CC   form a stable complex in soln. See also AAQ38442. (Updated on 25-MAR-2003
CC   to correct PN field.)
XX
SQ   Sequence 21 BP; 5 A; 3 C; 5 G; 8 T; 0 U; 0 Other;
```

```
  Query Match             100.0%;  Score 20;  DB 2;  Length 21;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db          2 TTATAAGGGTCGATGTCCAT 21


RESULT 4
AAZ98697
ID   AAZ98697 standard; DNA; 21 BP.
XX
AC   AAZ98697;
XX
DT   06-AUG-2003  (revised)
DT   20-JUN-2000  (first entry)
XX
DE   Human hepatitis B virus antisense oligonucleotide sequence.
XX
KW   Cytostatic; virucide; hepatotrophic; anti-inflammatory; antisense;
KW   hepatitis B virus; oncoprotein expression inhibitor; ss;
KW   asialoglycoprotein receptor.
XX
OS   Hepatitis B virus.
XX
PN   US6030954-A.
XX
PD   29-FEB-2000.
XX
PF   02-JUN-1995;   95US-00458633.
XX
PR   05-SEP-1991;   91US-00755083.
PR   04-NOV-1991;   91US-00788119.
PR   03-APR-1992;   92US-00864003.
PR   04-SEP-1992;   92US-00941368.
XX
PA   (UYCO-) UNIV CONNECTICUT.
XX
PI   Wu CH,  Wu GY;
XX
DR   WPI; 2000-223192/19.
XX
PT   Antisense oligonucleotides targeted to asialoglycoprotein receptor-
PT   bearing cells useful for inhibiting viral and oncoprotein RNA expression.
XX
PS   Example 1; Col 5; 12pp; English.
XX
CC   This sequence represents a human hepatitis B antisense oligonucleotide
CC   which can be used as a component of the soluble molecular complex of the
CC   invention. The invention relates to a soluble molecular complex
CC   comprising a single stranded antisense oligonucleotide which hybridises
CC   to an RNA in a target cell. The antisense oligonucleotide is complexed
CC   with a carrier comprised of a ligand for the asialoglycoprotein receptor
CC   and a polycation. The molecular complex has cytostatic, virucide,
CC   hepatotrophic and anti-inflammatory activity. The complex works through
CC   cell specific antisense inhibition of RNA expression. The molecular
CC   complex is used for inhibiting oncogene and viral (especially hepatitis)
CC   RNA expression in asialoglycoprotein receptor-bearing cells. (Updated on
CC   06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 21 BP; 5 A; 3 C; 5 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 3;  Length 21;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db          2 TTATAAGGGTCGATGTCCAT 21


RESULT 5
AAF58672
ID   AAF58672 standard; DNA; 21 BP.
XX
AC   AAF58672;
XX
DT   27-APR-2001  (first entry)
XX
DE   Hepatitis B virus HB probe.
XX
KW   Hepatitis B virus; HBV; antiinflammatory; immunosuppressive;
KW   hepatotropic; virucide; vaccine; Tupaia belangeri;
KW   immune-related disorder; transplantation rejection;
KW   selective immune down regulation; SIDR; probe; ss.
XX
OS   Hepatitis B virus.
XX
PN   EP1072271-A2.
XX
PD   31-JAN-2001.
XX
PF   17-JUL-2000; 2000EP-00115423.
XX
PR   16-JUL-1999;   99US-00356294.
XX
PA   (ENZO-) ENZO THERAPEUTICS INC.
XX
PI   Rabbani E,  Ilan Y,  Roy-Chowdhury J,  Engelhardt DL;
XX
DR   WPI; 2001-170934/18.
XX
PT   Native or non-native antigens, used for establishing selective immune
PT   down regulation, for transplantation, for treating or preventing
PT   undesirable immune reactions of vaccination, and for treating immune
PT   disorders.
XX
PS   Example 7; Page 18; 47pp; English.
XX
CC   The present sequence is a probe which was used in an example
CC   demonstrating disease symptoms induced in a small animal model, Tupaia
CC   belangeri, after infection by Hepatitis B virus (HBV). This example is
CC   provided in a specification relating to the use of native or non-native
CC   antigen or antigens, or their immunological equivalent, for preparing a
CC   pharmaceutical composition for use in transplantation processes, for
CC   treating or preventing undesirable immunological consequences of
CC   vaccination or immunisation in a subject, or for treating immune-related
CC   disorders. The invention provide unique selective immune down regulation
CC   (SIDR) applications in transplantation processes. They may be used for
CC   preventing or treating graft versus host rejection and for treating
CC   Crohn's disease, primary sclerosing cholangitis disease, primary biliary
CC   cirrhosis disease, primary Celliac's disease, primary autoimmune chronic
CC   active hepatitis, chronic liver rejection disease, immune-mediated liver
CC   fibrosis disease, immune-mediated vascular disorder, and immune-mediated
CC   muscle disorders affecting smooth muscle, striated muscle and blood
CC   vessel muscle
XX
SQ   Sequence 21 BP; 5 A; 3 C; 5 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 5;  Length 21;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db          2 TTATAAGGGTCGATGTCCAT 21


RESULT 6
ADB68574
ID   ADB68574 standard; DNA; 21 BP.
XX
AC   ADB68574;
XX
DT   04-DEC-2003  (first entry)
```

```
XX
DE   NG2 A-L-P conjugate DNA component used to target HBV c-gene.
XX
KW   homogeneous A-L-P conjugate; hepatic; chronic viral hepatitis; cirrhosis;
KW   malaria; viral infection; protozoan; cancer; hepatocellular carcinoma;
KW   HCC; ss; NG2; HBV; c-gene; core.
XX
OS   Hepatitis B virus.
XX
FH   Key             Location/Qualifiers
FT   modified_base    1. .21
FT                    /*tag=  b
FT                    /mod_base= OTHER
FT                    /note= "OTHER = phosphorothioate backbone"
FT   modified_base    1
FT                    /*tag=  a
FT                    /mod_base= OTHER
FT                    /note= "OTHER = Optionally linked to YEE(ahGalNAc)3-SMCC
FT                    and various chemical groups as shown in figures"
FT   modified_base    21
FT                    /*tag=  c
FT                    /mod_base= OTHER
FT                    /note= "OTHER = Optionally linked to chemical group as
FT                    shown in figure 5"
XX
PN   WO2003067209-A2.
XX
PD   14-AUG-2003.
XX
PF   21-JUN-2002; 2002WO-US019908.
XX
PR   22-JUN-2001; 2001US-00888164.
XX
PA   (CELL-) CELL WORKS INC.
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Ts'o POP,  Duff R,  Zhou Y,  Deamond S,  Roby C;
XX
DR   WPI; 2003-697456/66.
XX
PT   New homogeneous prodrug conjugate containing hepatic ligand for delivery
PT   of pathogen-specific oligomer useful for treating liver infections or
PT   cancer.
XX
PS   Claim 7; Page 83; 107pp; English.
XX
CC   The invention relates to a novel homogeneous conjugate comprising a
CC   hepatic ligand, bifunctional linker and biologically stable oligomer that
CC   binds to a sequence in a hepatic virus or pathogen and is released from
CC   the conjugate by hydrolysis or reduction. The conjugate of the invention
CC   may be useful during the treatment of liver diseases including chronic
CC   viral hepatitis, cirrhosis, malaria, viral or protozoan infection and
CC   cancer, such as hepatocellular carcinoma (HCC). The current sequence is
CC   that of the NG2 A-L-P conjugate DNA component of the invention which was
CC   used to target the Hepatitis B virus (HBV) c (core)-gene.
XX
SQ   Sequence 21 BP; 5 A; 3 C; 5 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 10;  Length 21;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db          2 TTATAAGGGTCGATGTCCAT 21


RESULT 7
ADB68572
ID   ADB68572 standard; RNA; 21 BP.
XX
AC   ADB68572;
XX
DT   04-DEC-2003  (first entry)
XX
DE   A-L-P conjugate-related RNA oligomer 3.
XX
KW   homogeneous A-L-P conjugate; hepatic; chronic viral hepatitis; cirrhosis;
KW   malaria; viral infection; protozoan; cancer; hepatocellular carcinoma;
KW   HCC; ss.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   modified_base    1. .21
FT                    /*tag=  a
FT                    /mod_base= OTHER
FT                    /note= "OTHER = 2'-O-methyl ribose alternating methyl-
FT                    phosphonate-phosphodiester backbone"
XX
PN   WO2003067209-A2.
XX
PD   14-AUG-2003.
XX
PF   21-JUN-2002; 2002WO-US019908.
XX
PR   22-JUN-2001; 2001US-00888164.
XX
PA   (CELL-) CELL WORKS INC.
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Ts'o POP,  Duff R,  Zhou Y,  Deamond S,  Roby C;
XX
DR   WPI; 2003-697456/66.
XX
PT   New homogeneous prodrug conjugate containing hepatic ligand for delivery
PT   of pathogen-specific oligomer useful for treating liver infections or
PT   cancer.
XX
PS   Example 2; Page 40; 107pp; English.
XX
CC   The invention relates to a novel homogeneous conjugate comprising a
CC   hepatic ligand, bifunctional linker and biologically stable oligomer that
CC   binds to a sequence in a hepatic virus or pathogen and is released from
CC   the conjugate by hydrolysis or reduction. The conjugate of the invention
CC   may be useful during the treatment of liver diseases including chronic
CC   viral hepatitis, cirrhosis, malaria, viral or protozoan infection and
CC   cancer, such as hepatocellular carcinoma (HCC). The current sequence is
CC   that of the A-L-P conjugate-related RNA oligomer 3 of the invention.
XX
SQ   Sequence 21 BP; 5 A; 3 C; 5 G; 0 T; 8 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 10;  Length 21;
  Best Local Similarity   80.0%;  Pred. No. 1.2;
  Matches   16;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ::|:|||||:||||||||||
Db          2 UUAUAAGGGUCGAUGUCCAU 21


RESULT 8
AAT72577
ID   AAT72577 standard; DNA; 27 BP.
XX
AC   AAT72577;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV94b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
```

```
FT   misc_feature    1. .27
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV94b which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 27 BP; 8 A; 4 C; 5 G; 10 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 27;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db          8 TTATAAGGGTCGATGTCCAT 27


RESULT 9
AAQ80499/c
ID   AAQ80499 standard; DNA; 28 BP.
XX
AC   AAQ80499;
XX
DT   25-MAR-2003  (revised)
DT   23-AUG-1995  (first entry)
XX
DE   Primer to amplify hepatitis B virus core region.
XX
KW   hepatitis B virus; X region; core region; primer; PCR; amplification;
KW   polymerase chain reaction; detection; viral infection; ss.
XX
OS   Synthetic.
XX
PN   WO9429483-A1.
XX
PD   22-DEC-1994.
XX
PF   03-JUN-1994;   94WO-US006360.
XX
PR   08-JUN-1993;   93US-00074346.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Feitelson M,  Duan L,  Guo J;
XX
DR   WPI; 1995-036505/05.
XX
PT   Detection of hepatitis B virus (HBV) variants having deletions in the X
PT   region - by detection of antibodies against HBV polymerase and HB X
PT   antigen.
XX
PS   Claim 3; Page 34; 45pp; English.
XX
CC   This primer designated MF03 covers nucleotide bases 1903-1949 at the
CC   beginning of the hepatitis B virus (HBV) core open reading frame. It is
CC   used with MFO4 (AAQ80500) to amplify the core gene. The primers allow the
CC   detection of a specific class of HBV variants. They are useful for
CC   demonstrating the presence of productive virus infection and may prove
CC   useful in monitoring therapeutics. (Updated on 25-MAR-2003 to correct PN
CC   field.)
XX
SQ   Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db         20 TTATAAGGGTCGATGTCCAT 1


RESULT 10
AAV45201/c
ID   AAV45201 standard; DNA; 28 BP.
XX
AC   AAV45201;
XX
DT   19-OCT-1998  (first entry)
XX
DE   Primer MFO3.
XX
KW   ss; PCR; primer; amplification; viral infection; bacterial infection;
KW   immune response; hepatitis B virus.
XX
OS   Mus sp.
XX
PN   WO9829121-A1.
XX
PD   09-JUL-1998.
XX
PF   02-JAN-1998;   98WO-US004116.
XX
PR   02-JAN-1997;   97US-0034596P.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Michaels F,  Block T;
XX
DR   WPI; 1998-387782/33.
XX
PT   Modulating immune responses in mammals infected with infectious agent(s)
PT   - e.g. to reduce pathogenicity caused by immune responses in cases where
PT   the infectious agent has limited pathogenicity.
XX
PS   Example 2; Page 37; 55pp; English.
XX
CC   The primers AAV45201 and AAV45202 were used to detect the presence of a
CC   HBV genome which had been microinjected into embryos of SCID mice in an
CC   example to demonstrate modulating an immune response in a mammal infected
CC   with an infectious agent. This comprises transmucosal administration of a
CC   composition comprising an epitope which is located in close proximity to
CC   the immune response. The process may be used in treatment of mammals
CC   which are acutely or chronically infected with infectious agents, such as
CC   viruses or bacteria. It may be used to increase the immune response, or
CC   it may be used to decrease the immune response in cases where the
CC   infectious agent itself exhibits limited pathogenicity but the immune
CC   response to the infectious agent causes more significant pathogenicity.
CC   This can be the case in, e.g. hepatitis B virus (HBV) infection. The
```

```
CC   process can modulate undesirable autoimmune responses exhibited by
CC   mammals infected with viral, bacterial and parasitic agents. It can
CC   prevent life-long disabilities which result from these infections
XX
SQ   Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db         20 TTATAAGGGTCGATGTCCAT 1


RESULT 11
AAA62585/c
ID   AAA62585 standard; DNA; 28 BP.
XX
AC   AAA62585;
XX
DT   22-NOV-2000  (first entry)
XX
DE   Transgenic SCID mouse hepatitis virus transgene PCR primer MF03.
XX
KW   Mouse; SCID; severe combined immunodeficiency; transgenic mouse;
KW   hepatitis virus; hepatitis B; hepatitis C; chronic liver disease;
KW   PCR primer; ss.
XX
OS   Mus sp.
XX
PN   US6087556-A.
XX
PD   11-JUL-2000.
XX
PF   07-JAN-1998;   98US-00003200.
XX
PR   02-MAY-1996;   96US-00641803.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Siracusa L,  Feitelson M;
XX
DR   WPI; 2000-523731/47.
XX
PT   Transgenic mouse useful in methods for evaluating interactions of
PT   chemical, drug or immunomodulating agent with hepatitis virus, lacks
PT   functional T-cells and B-cells and is capable of replicating hepatitis
PT   viruses.
XX
PS   Example 3; Col 8; 7pp; English.
XX
CC   The present sequence is a PCR primer used to detect a transgene in severe
CC   combined immunodeficient (SCID) mice. Transgenic immunodeficient mice
CC   were produced that are not tolerant to hepatitis viral antigens, lack
CC   functional T-cells and B-cells and contain integrated hepatitis virus DNA
CC   in the somatic and germ cells. The hepatitis virus gene is expressed and
CC   the hepatitis virus is replicated in the transgenic mouse. The mouse may
CC   be used as an animal model for evaluating interactions of a chemical,
CC   drug or immunomodulating agent with a hepatitis virus. It is also useful
CC   for the assessment of anti-viral and immunomodulatory intervention
CC   therapies, including the screening of drug candidates. It can be used to
CC   analyse the virus-host relationship, to evaluate the relationship between
CC   the virus and chemicals metabolised and/or detoxified by the liver, and
CC   to identify cellular biochemical pathways contributing to the development
CC   and progression of chronic liver disease. The transgenic mouse is thus
CC   useful for elucidating the effects of hepatitis virus on hepatic
CC   metabolism
XX
SQ   Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 3;  Length 28;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db         20 TTATAAGGGTCGATGTCCAT 1


RESULT 12
ADN36070/c
ID   ADN36070 standard; DNA; 28 BP.
XX
AC   ADN36070;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Probe #151 to determine effect of long term lamivudine treatment of HBV.
XX
KW   ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS   Hepatitis B virus.
XX
PN   WO2004031729-A2.
XX
PD   15-APR-2004.
XX
PF   01-OCT-2003; 2003WO-US031121.
XX
PR   01-OCT-2002; 2002US-0415301P.
XX
PA   (GEOU ) UNIV GEORGETOWN.
XX
PI   Korba BE,  Ciancio A,  Gerin JL;
XX
DR   WPI; 2004-348004/32.
XX
PT   Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT   3TC therapy comprises determining whether the HBV bears a nucleic acid
PT   encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS   Claim 31; SEQ ID NO 151; 107pp; English.
XX
CC   The invention relates to a method of predicting the long term response of
CC   a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC   the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC   leucine at amino acid position (aa) 91 in the DNA polymerase region
CC   (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC   604) in the DNA polymerase region of HBV. The method comprises
CC   determining whether the HBV carried by the host bears one or more of the
CC   following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC   codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC   T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC   promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC   changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC   and C1962A representing specific double point mutations in the DNA
CC   precore/core promoter or ORF region. The method and kit is useful in
CC   predicting the long-term response of a host of HBV to 3TC therapy (also
CC   known as lamivudine). This sequence represents an oligonucleotide
CC   sequence used in the method of the invention to detect a mutation in the
CC   above mentioned sequences.
XX
SQ   Sequence 28 BP; 11 A; 6 C; 5 G; 6 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 12;  Length 28;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db         23 TTATAAGGGTCGATGTCCAT 4


RESULT 13
ADN36068/c
```

```
ID   ADN36068 standard; DNA; 28 BP.
XX
AC   ADN36068;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Probe #149 to determine effect of long term lamivudine treatment of HBV.
XX
KW   ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS   Hepatitis B virus.
XX
PN   WO2004031729-A2.
XX
PD   15-APR-2004.
XX
PF   01-OCT-2003; 2003WO-US031121.
XX
PR   01-OCT-2002; 2002US-0415301P.
XX
PA   (GEOU ) UNIV GEORGETOWN.
XX
PI   Korba BE,  Ciancio A,  Gerin JL;
XX
DR   WPI; 2004-348004/32.
XX
PT   Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT   3TC therapy comprises determining whether the HBV bears a nucleic acid
PT   encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS   Claim 31; SEQ ID NO 149; 107pp; English.
XX
CC   The invention relates to a method of predicting the long term response of
CC   a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC   the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC   leucine at amino acid position (aa) 91 in the DNA polymerase region
CC   (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC   604) in the DNA polymerase region of HBV. The method comprises
CC   determining whether the HBV carried by the host bears one or more of the
CC   following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC   codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC   T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC   promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC   changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC   and C1962A representing specific double point mutations in the DNA
CC   precore/core promoter or ORF region. The method and kit is useful in
CC   predicting the long-term response of a host of HBV to 3TC therapy (also
CC   known as lamivudine). This sequence represents an oligonucleotide
CC   sequence used in the method of the invention to detect a mutation in the
CC   above mentioned sequences.
XX
SQ   Sequence 28 BP; 11 A; 6 C; 6 G; 5 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 12;  Length 28;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||||:|:|||:
Db         25 TTATAAGGGTCGATGTCCAT 6


RESULT 14
ADN36071/c
ID   ADN36071 standard; DNA; 28 BP.
XX
AC   ADN36071;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Probe #152 to determine effect of long term lamivudine treatment of HBV.
XX
KW   ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS   Hepatitis B virus.
XX
PN   WO2004031729-A2.
XX
PD   15-APR-2004.
XX
PF   01-OCT-2003; 2003WO-US031121.
XX
PR   01-OCT-2002; 2002US-0415301P.
XX
PA   (GEOU ) UNIV GEORGETOWN.
XX
PI   Korba BE,  Ciancio A,  Gerin JL;
XX
DR   WPI; 2004-348004/32.
XX
PT   Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT   3TC therapy comprises determining whether the HBV bears a nucleic acid
PT   encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS   Claim 31; SEQ ID NO 152; 107pp; English.
XX
CC   The invention relates to a method of predicting the long term response of
CC   a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC   the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC   leucine at amino acid position (aa) 91 in the DNA polymerase region
CC   (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC   604) in the DNA polymerase region of HBV. The method comprises
CC   determining whether the HBV carried by the host bears one or more of the
CC   following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC   codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC   T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC   promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC   changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC   and C1962A representing specific double point mutations in the DNA
CC   precore/core promoter or ORF region. The method and kit is useful in
CC   predicting the long-term response of a host of HBV to 3TC therapy (also
CC   known as lamivudine). This sequence represents an oligonucleotide
CC   sequence used in the method of the invention to detect a mutation in the
CC   above mentioned sequences.
XX
SQ   Sequence 28 BP; 11 A; 6 C; 4 G; 7 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 20;  DB 12;  Length 28;
  Best Local Similarity   85.0%;  Pred. No. 1.2;
  Matches   17;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||||:|:|||:
Db         22 TTATAAGGGTCGATGTCCAT 3


RESULT 15
ADN36073/c
ID   ADN36073 standard; DNA; 28 BP.
XX
AC   ADN36073;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Probe #154 to determine effect of long term lamivudine treatment of HBV.
XX
KW   ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS   Hepatitis B virus.
XX
PN   WO2004031729-A2.
XX
PD   15-APR-2004.
XX
PF   01-OCT-2003; 2003WO-US031121.
XX
```

PR 01-OCT-2002; 2002US-0415301P.
XX
PA (GEOU) UNIV GEORGETOWN.
XX
PI Korba BE, Ciancio A, Gerin JL;
XX
DR WPI; 2004-348004/32.
XX
PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS Claim 31; SEQ ID NO 154; 107pp; English.
XX
CC The invention relates to a method of predicting the long term response of
CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC leucine at amino acid position (aa) 91 in the DNA polymerase region
CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC 604) in the DNA polymerase region of HBV. The method comprises
CC determining whether the HBV carried by the host bears one or more of the
CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC and C1962A representing specific double point mutations in the DNA
CC precore/core promoter or ORF region. The method and kit is useful in
CC predicting the long-term response of a host of HBV to 3TC therapy (also
CC known as lamivudine). This sequence represents an oligonucleotide
CC sequence used in the method of the invention to detect a mutation in the
CC above mentioned sequences.
XX
SQ Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 85.0%; Pred. No. 1.2;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||||||:|:|||:
Db         20 TTATAAGGGTCGATGTCCAT 1
```

Search completed: March 17, 2005, 06:48:44
Job time : 172.333 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-64
Perfect score: 20
Sequence: 1 ttataagggtcgauguccau 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	18.4	92.0	706	4	BI888237	BI888237 ZF637-1-0
	2	17.4	87.0	738	8	AZ195485	AZ195485 SP_1030_A
	3	16.8	84.0	209	4	BM024704	BM024704 fu71e06.x
	4	16.8	84.0	211	6	CD014361	CD014361 hac31a04.
c	5	16.8	84.0	331	9	CR405240	CR405240 Arabidops
	6	16.8	84.0	420	7	H98675	H98675 yx17d02.s1
	7	16.8	84.0	426	4	BM571757	BM571757 fx05f04.x
	8	16.8	84.0	433	4	BI840986	BI840986 fq41f02.x
c	9	16.8	84.0	438	2	AW147602	AW147602 da13f06.y
	10	16.8	84.0	475	7	H98688	H98688 yx17h02.s1
c	11	16.8	84.0	477	4	BM571990	BM571990 fx05f04.y
c	12	16.8	84.0	495	5	BW093445	BW093445 BW093445
	13	16.8	84.0	508	4	BM072332	BM072332 fv08d10.x
	14	16.8	84.0	528	5	BW234048	BW234048 BW234048
c	15	16.8	84.0	551	5	BW392185	BW392185 BW392185
	16	16.8	84.0	553	4	BM185735	BM185735 fv71e09.x
	17	16.8	84.0	562	4	BM185703	BM185703 fv71a05.x
	18	16.8	84.0	563	4	BM005095	BM005095 fu64c11.x
	19	16.8	84.0	580	4	BI983042	BI983042 fu42f10.x
	20	16.8	84.0	584	4	BM005060	BM005060 fu63g11.x
	21	16.8	84.0	586	4	BM024740	BM024740 fu72b11.x
	22	16.8	84.0	586	9	CR333657	CR333657 Medicago
	23	16.8	84.0	600	5	BW338052	BW338052 BW338052
	24	16.8	84.0	621	6	CD014379	CD014379 hac31c04.
	25	16.8	84.0	695	6	CD237681	CD237681 FNPAOC01
c	26	16.8	84.0	710	5	BW018497	BW018497 BW018497
c	27	16.8	84.0	711	9	BX222138	BX222138 Danio rer
	28	16.8	84.0	714	8	AQ648647	AQ648647 RPCI93-Ec
c	29	16.8	84.0	729	4	BJ705222	BJ705222 BJ705222
	30	16.8	84.0	733	1	AI635930	AI635930 tz82c11.x
c	31	16.8	84.0	734	9	BX138600	BX138600 Danio rer
c	32	16.8	84.0	738	9	BX245774	BX245774 Danio rer
c	33	16.8	84.0	774	5	BW025653	BW025653 BW025653
c	34	16.8	84.0	775	9	BX170776	BX170776 Danio rer
c	35	16.8	84.0	780	9	BX144130	BX144130 Danio rer
c	36	16.8	84.0	796	5	BW023432	BW023432 BW023432
c	37	16.8	84.0	821	7	CK129749	CK129749 AGENCOURT
c	38	16.8	84.0	860	9	CG002233	CG002233 ZUABD83TH
	39	16.4	82.0	413	5	BP569691	BP569691 BP569691
c	40	16.4	82.0	492	9	CG596834	CG596834 OST259526
c	41	16.4	82.0	573	9	CG594911	CG594911 OST253619
	42	16.4	82.0	584	7	CK529842	CK529842 rswfa0_01
	43	16.4	82.0	597	7	CK529191	CK529191 rswfa0_00
	44	16.4	82.0	623	7	CK528881	CK528881 rswfa0_00
	45	16.4	82.0	656	5	BW362244	BW362244 BW362244

ALIGNMENTS

RESULT 1
BI888237/c
LOCUS BI888237 706 bp mRNA linear EST 12-OCT-2001
DEFINITION ZF637-1-002492 Zebrafish shield stage whole embryo cDNA library MPMGp637 Danio rerio cDNA clone MPMGp637_18P2;MPMGp637P0218 5', mRNA sequence.
ACCESSION BI888237
VERSION BI888237.1 GI:16095508
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 706)
AUTHORS Clark,M., Aanstad,P., Hennig,S., Johnson,S.L. and Lehrach,H.
TITLE EST sequencing of a zebrafish shield stage cDNA library normalised by oligonucleotide fingerprinting
JOURNAL Unpublished (2001)
COMMENT Contact: Hennig S
laboraty 123, dept.Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Ihnestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
5' EST sequencing of clones from a zebrafish shield stage library, normalised from 55,000 starting clones by oligonucleotide fingerprinting
High quality sequence stop: 706.
FEATURES Location/Qualifiers
source 1. .706
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MPMGp637_18P2;MPMGp637P0218"
/tissue_type="whole embryo"
/dev_stage="shield stage, 6 hrs post-fertilisation"
/lab_host="E.coli, XL1 blue MRF"
/clone_lib="Zebrafish shield stage whole embryo cDNA library MPMGp637"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI; oligo-dT-NotI primed, SalI adaptors, directionally cloned, library normalised by oligonucleotide fingerprinting"
ORIGIN

Query Match 92.0%; Score 18.4; DB 4; Length 706;

```
  Best Local Similarity   80.0%;  Pred. No. 55;
  Matches   16;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||||||||||||:|: ||:
Db        499 TTATAAGGGTCGATGTGCAT 480


RESULT 2
AZ195485
LOCUS       AZ195485                738 bp    DNA     linear   GSS 30-AUG-2000
DEFINITION  SP_1030_A1_H09_SP6E Strongylocentrotus purpuratus, purple sea
            urchin, sperm genomic BAC library Strongylocentrotus purpuratus
            genomic clone Plate=1030 Col=17 Row=O, genomic survey sequence.
ACCESSION   AZ195485
VERSION     AZ195485.1  GI:8378664
KEYWORDS    GSS.
SOURCE      Strongylocentrotus purpuratus
  ORGANISM  Strongylocentrotus purpuratus
            Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
            Echinoidea; Euechinoidea; Echinacea; Echinoida;
            Strongylocentrotidae; Strongylocentrotus.
REFERENCE   1  (bases 1 to 738)
  AUTHORS   Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
            Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T.,
            Wray,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J, Davidson,E.H.
            and Hood,L.
  TITLE     A sea urchin genome project: Sequence scan, virtual map, and
            additional resources
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)
  MEDLINE   20402566
   PUBMED   10920195
COMMENT     Contact: Cameron, RA, Davidson, EH, Hood, L
            Division of Biology 156-29
            California Institute of Technology
            Pasadena California 91125, USA
            Tel: (626) 395-8421
            Fax: (626) 793-3047
            Email: acameron@caltech.edu
            Plate: 1030  row: O  column: 17
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1. .738
                     /organism="Strongylocentrotus purpuratus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7668"
                     /clone="Plate=1030 Col=17 Row=O"
                     /clone_lib="Strongylocentrotus purpuratus, purple sea
                     urchin, sperm genomic BAC library"
                     /note="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli
                     DH10B"
ORIGIN

  Query Match            87.0%;  Score 17.4;  DB 8;  Length 738;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+02;
  Matches   15;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TATAAGGGTCGAUGUCCAU 20
              |||||||| |||:|:|||:
Db          9 TATAAGGGGCGATGTCCAT 27


RESULT 3
BM024704
LOCUS       BM024704                209 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  fu71e06.x1 zebrafish adult brain Danio rerio cDNA clone
            IMAGE:5334995 3', mRNA sequence.
ACCESSION   BM024704
VERSION     BM024704.1  GI:16539060
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 209)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
            Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Seq primer: -40UP
            High quality sequence stop: 200.
FEATURES             Location/Qualifiers
     source          1. .209
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:5334995"
                     /sex="mixed male and female"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /lab_host="E. coli DH10B"
                     /clone_lib="zebrafish adult brain"
                     /note="Vector: pZIPLOX; Site_1: NotI; Site_2: SalI;
                     Original library was constructed in lambdaZIPLOX.  Mass
                     excision of the cDNA library was performed to yield
                     pZIPLOX plasmids. Insert check was done in original
                     library."
ORIGIN

  Query Match            84.0%;  Score 16.8;  DB 4;  Length 209;
  Best Local Similarity   75.0%;  Pred. No. 3.5e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||| ||||||||:|: ||:
Db        176 TTATTAGGGTCGATGTGCAT 195


RESULT 4
CD014361
LOCUS       CD014361                211 bp    mRNA    linear   EST 02-MAY-2003
DEFINITION  hac31a04.x1 MPZFRiken1 Danio rerio cDNA clone IMAGE:6923625 3',
            mRNA sequence.
ACCESSION   CD014361
VERSION     CD014361.1  GI:30330820
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 211)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
            Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
```

```
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            Oligo-dT primed cDNA was produced from intestine (two independent
            samples) and liver using
            5'-GAGAGAGAGAAGGATCCAAXXXXXXTTTTTTTTTTTTTTTTVN-3' with the
            following tags substituting for XXXXXX: CAAGAG (tag C04): liver;
            CGGTAT (tag C12): intestine; and CGTATG (tag D01): intestine. cDNA
            was pooled and size- selected for a 1.6 kb average insert. Library
            was constructed using the Cap- Trapper method as described in
            Genomics 2001: 77(1-2)79-90. Library donated by M. Pack, M.D.
            (University of Pennsylvania School of Medicine).
            Seq primer: T3 from Gibco
            High quality sequence stop: 211.
FEATURES             Location/Qualifiers
     source          1. .211
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:6923625"
                     /tissue_type="intestine (2 samples) and liver (pooled)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="MPZFRikenl"
                     /note="Vector: pBluescriptR; Site_1: XhoI; Site_2: BamHI;
                     Oligo-dT primed cDNA was produced from intestine (two
                     independent samples) and liver using
                     5'-GAGAGAGAGAAGGATCCAAXXXXXXTTTTTTTTTTTTTTTTVN-3' with the
                     following tags substituting for XXXXXX: CAAGAG (tag C04):
                     liver; CGGTAT (tag C12): intestine; and CGTATG (tag D01):
                     intestine. cDNA was pooled and size- selected for a 1.6 kb
                     average insert. Library was constructed using the Cap-
                     Trapper method as described in Genomics 2001:
                     77(1-2)79-90. Library donated by M. Pack, M.D. (University
                     of Pennsylvania School of Medicine)."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 6;  Length 211;
  Best Local Similarity   75.0%;  Pred. No. 3.5e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||| ||||||||:|: ||:
Db         32 TTATTAGGGTCGATGTGCAT 51


RESULT 5
CR405240/c
LOCUS       CR405240                331 bp    DNA     linear   GSS 02-MAY-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-876H09-026468,
            genomic survey sequence.
ACCESSION   CR405240
VERSION     CR405240.1  GI:46945968
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
  TITLE     GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
  JOURNAL   Bioinformatics 19 (11), 1441-1442 (2003)
  MEDLINE   22755829
   PUBMED   12874060
REFERENCE   2
  AUTHORS   Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
            Weisshaar,B.
  TITLE     An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
            flanking sequence tag-based reverse genetics
  JOURNAL   Plant Mol. Biol. 53 (1-2), 247-259 (2003)
  MEDLINE   23117147
   PUBMED   14756321
REFERENCE   3
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
            Weisshaar,B.
  TITLE     High-throughput generation of sequence indexes from T-DNA
            mutagenized Arabidopsis thaliana lines
  JOURNAL   BioTechniques 35 (6), 1164-1168 (2003)
   PUBMED   14682050
REFERENCE   4  (bases 1 to 331)
  AUTHORS   Rosso,M.G., Strizhov,N., Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2004) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence has been recovered from the left border of the T-DNA.
            It indicates an insertion within the locus defined by BAC clone
            F9D12. Details on the protocols used for generation of the sequence
            are described in References 1-3. The sequences are generated at the
            MPI for Plant Breeding Research in the context of the GABI-Kat
            project. GABI-Kat is part of the German Plant Genomics program
            designated 'GABI'. Information on line availability can be found
            at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .331
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-876H09-026468"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 331;
  Best Local Similarity   75.0%;  Pred. No. 3.7e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||||| ||:|:|||:
Db         35 TTATAAGGGAGGATGTCCAT 16


RESULT 6
H98675
LOCUS       H98675                  420 bp    mRNA    linear   EST 15-DEC-1995
DEFINITION  yx17d02.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
            IMAGE:261987 3', mRNA sequence.
ACCESSION   H98675
VERSION     H98675.1  GI:1123343
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 420)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
```

```
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 237
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: m13 -40 forward
            High quality sequence stop: 237.
FEATURES             Location/Qualifiers
     source          1. .420
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3871629"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:261987"
                     /sex="Male"
                     /tissue_type="melanocyte"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares melanocyte 2NbHM"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. RNA from normal foreskin melanocytes
                     (FS374) was kindly provided by Dr. Anthony P. Albino."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 7;  Length 420;
  Best Local Similarity   75.0%;  Pred. No. 3.8e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              || | ||||||||:|:|||:
Db         94 TTCTCAGGGTCGATGTCCAT 113


RESULT 7
BM571757
LOCUS       BM571757                 426 bp    mRNA    linear   EST 21-FEB-2002
DEFINITION  fx05f04.x1 Gong zebrafish ovary Danio rerio cDNA clone
            IMAGE:5619198 3', mRNA sequence.
ACCESSION   BM571757
VERSION     BM571757.1  GI:18853740
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
            Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            The library was constructed by Dr. Z. Gong. DNA Sequencing by:
            Washington University Genome Sequencing Center St. Louis. Please
            contact Zhiyuan Gong for further information on this library
            (National University of Singapore, Department of Biological
            Sciences, Lower Kent Ridge Road, Singapore 119260).
            Seq primer: T7 from Gibco.
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:5619198"
                     /sex="female"
                     /dev_stage="4-5 month"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Gong zebrafish ovary"
                     /note="Organ: ovary (pooled); Vector: pBluescript SK-;
                     Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolatd from
                     the ovaries of 2 female adult zebrafish (4-5 month  old).
                     cDNAs were made using oligo-dT primers and inserted into
                     lambda ZAP II vector  (Stratagene) by Dr. Z. Gong, in vivo
                     mass-excised to pBluescript SK- following the  Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact Zhiyuan Gong for further information on
                     this library (National University  of Singapore,
                     Department of Biological Sciences, Lower Kent Ridge Road,
                     Singapore 119260)."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 4;  Length 426;
  Best Local Similarity   75.0%;  Pred. No. 3.9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||| ||||||||:|: ||:
Db        205 TTATTAGGGTCGATGTGCAT 224


RESULT 8
BI840986
LOCUS       BI840986                 433 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  fq41f02.x1 zebrafish adult brain Danio rerio cDNA clone
            IMAGE:4833939 3', mRNA sequence.
ACCESSION   BI840986
VERSION     BI840986.1  GI:15953509
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 433)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
            Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
```

```
          info@genomesystems.com) and Research Genetics, Huntsville, Alabama
          (web address: www.resgen.com) (email contact: info@resgen.com) and
          RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
          www.rzpd.de)
          Seq primer: -40UP
          High quality sequence stop: 396.
FEATURES             Location/Qualifiers
     source          1. .433
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:4833939"
                     /sex="mixed male and female"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /lab_host="E. coli DH10B"
                     /clone_lib="zebrafish adult brain"
                     /note="Vector: pZIPLOX; Site_1: NotI; Site_2: SalI;
                     Original library was constructed in lambdaZIPLOX.  Mass
                     excision of the cDNA library was performed to yield
                     pZIPLOX plasmids. Insert check was done in original
                     library."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 4;  Length 433;
  Best Local Similarity   75.0%;  Pred. No. 3.9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||| |||||||:|: ||:
Db        182 TTATTAGGGTCGATGTGCAT 201


RESULT 9
AW147602/c
LOCUS       AW147602                438 bp    mRNA    linear   EST 19-FEB-2003
DEFINITION  da13f06.y1 normalized Xenopus laevis gastrula Xenopus laevis cDNA
            clone XENOPUS_SOURCE_ID:xlnga001c12 5' similar to SW:AOP2_HUMAN
            P30041 ANTIOXIDANT PROTEIN 2 ;, mRNA sequence.
ACCESSION   AW147602
VERSION     AW147602.1  GI:6195498
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L.,
            Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y.,
            Person,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Xenopus EST project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Sandy Clifton, Ph.D.
            WashU Xenopus EST project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by Bruce Blumberg
             Library normalized by Jihwan Song
             DNA Sequencing by: Washington University Genome Sequencing Center
            Seq primer: -40RP from Gibco
            High quality sequence stop: 385.
FEATURES             Location/Qualifiers
     source          1. .438
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="XENOPUS_SOURCE_ID:xlnga001c12"
                     /tissue_type="gastrula (stages 10.5, 11.5 mixed)"
                     /lab_host="Top-10 F'"
                     /clone_lib="normalized Xenopus laevis gastrula"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from 2ug of poly A+ RNA (equal
                     parts from stage 10.5 and stage 11.5 gastrulae).
                     EcoRI-XhoI cut cDNA was then ligated into UniZap-XR
                     (Stratagene) with EcoRI at the 5' end and XhoI at the 3'
                     end. SS-library phagemids were prepared by mass excision
                     from the original library and normalized by hybridization
                     to biotinylated driver (prepared from the same library by
                     PCR) to Cot-omega of 11. After removal of hybrids and
                     excess driver by streptavidin sepharose chromatography,
                     the ss-phagemids were made double stranded and
                     electroporated into Top-10 F'. Original library
                     contruction by Bruce Blumberg (Cho et al. 1991 Cell 67,
                     1111-1120). Normalized by Jihwan Song (Song, Cho and
                     Blumberg, unpublished). Note: This is a Xenopus Gene
                     Collection (XGC) library."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 2;  Length 438;
  Best Local Similarity   75.0%;  Pred. No. 3.9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||| || ||:|:|||:
Db        299 TTATAAGAGTTGATGTCCAT 280


RESULT 10
H98688
LOCUS       H98688                  475 bp    mRNA    linear   EST 15-DEC-1995
DEFINITION  yx17h02.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
            IMAGE:262035 3', mRNA sequence.
ACCESSION   H98688
VERSION     H98688.1  GI:1123356
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 475)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 324
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: m13 -40 forward
            High quality sequence stop: 324.
FEATURES             Location/Qualifiers
     source          1. .475
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3871677"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:262035"
                     /sex="Male"
                     /tissue_type="melanocyte"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares melanocyte 2NbHM"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
```

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."

ORIGIN

```
  Query Match             84.0%;  Score 16.8;  DB 7;  Length 475;
  Best Local Similarity   75.0%;  Pred. No. 3.9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              || | ||||||||:|:|||:
Db         93 TTCTCAGGGTCGATGTCCAT 112
```

RESULT 11
BM571990/c

```
LOCUS       BM571990                477 bp    mRNA    linear   EST 21-FEB-2002
DEFINITION  fx05f04.y1 Gong zebrafish ovary Danio rerio cDNA clone
            IMAGE:5619198 5', mRNA sequence.
ACCESSION   BM571990
VERSION     BM571990.1  GI:18853977
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
            Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Other_ESTs: fx05f04.x1
            Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            The library was constructed by Dr. Z. Gong. DNA Sequencing by:
            Washington University Genome Sequencing Center St. Louis. Please
            contact Zhiyuan Gong for further information on this library
            (National University of Singapore, Department of Biological
            Sciences, Lower Kent Ridge Road, Singapore 119260).
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 394.
FEATURES             Location/Qualifiers
     source          1. .477
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:5619198"
                     /sex="female"
                     /dev_stage="4-5 month"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Gong zebrafish ovary"
                     /note="Organ: ovary (pooled); Vector: pBluescript SK-;
                     Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolatd from
                     the ovaries of 2 female adult zebrafish (4-5 month old).
                     cDNAs were made using oligo-dT primers and inserted into
                     lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
                     mass-excised to pBluescript SK- following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact Zhiyuan Gong for further information on
                     this library (National University of Singapore,
                     Department of Biological Sciences, Lower Kent Ridge Road,
                     Singapore 119260)."
ORIGIN
```

```
  Query Match             84.0%;  Score 16.8;  DB 4;  Length 477;
  Best Local Similarity   75.0%;  Pred. No. 3.9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||| ||||||||:|: ||:
Db        291 TTATTAGGGTCGATGTGCAT 272
```

RESULT 12
BW093445/c

```
LOCUS       BW093445                495 bp    mRNA    linear   EST 23-OCT-2002
DEFINITION  BW093445 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
            intestinalis cDNA clone rcitb046h04 3', mRNA sequence.
ACCESSION   BW093445
VERSION     BW093445.1  GI:24269247
KEYWORDS    EST.
SOURCE      Ciona intestinalis
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
  TITLE     Expressed genes in Ciona intestinalis (2002c)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .495
                     /organism="Ciona intestinalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:7719"
                     /clone="rcitb046h04"
                     /tissue_type="whole animal"
                     /dev_stage="tailbud embryo"
                     /clone_lib="Nori Satoh unpublished cDNA library, tailbud
                     embryo"
ORIGIN
```

```
  Query Match             84.0%;  Score 16.8;  DB 5;  Length 495;
  Best Local Similarity   75.0%;  Pred. No. 3.9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||| || ||:|:|||:
Db        346 TTATAAGTGTGGATGTCCAT 327
```

RESULT 13
BM072332

```
LOCUS       BM072332                508 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  fv08d10.x1 zebrafish adult brain Danio rerio cDNA clone
            IMAGE:5386027 3', mRNA sequence.
ACCESSION   BM072332
VERSION     BM072332.1  GI:16915252
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
```

```
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
            Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Seq primer: -40UP
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .508
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:5386027"
                     /sex="mixed male and female"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /lab_host="E. coli DH10B"
                     /clone_lib="zebrafish adult brain"
                     /note="Vector: pZIPLOX; Site_1: NotI; Site_2: SalI;
                     Original library was constructed in lambdaZIPLOX.  Mass
                     excision of the cDNA library was performed to yield
                     pZIPLOX plasmids. Insert check was done in original
                     library."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 4;  Length 508;
  Best Local Similarity   75.0%;  Pred. No. 3.9e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              |||| ||||||||:|: ||:
Db        180 TTATTAGGGTCGATGTGCAT 199


RESULT 14
BW234048
LOCUS       BW234048                528 bp    mRNA    linear   EST 07-NOV-2002
DEFINITION  BW234048 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
            intestinalis cDNA clone citb046h04 5', mRNA sequence.
ACCESSION   BW234048
VERSION     BW234048.1  GI:24755889
KEYWORDS    EST.
SOURCE      Ciona intestinalis
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
  TITLE     Expressed genes in Ciona intestinalis (2002c)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Ciona intestinalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:7719"
                     /clone="citb046h04"
                     /tissue_type="whole animal"
                     /dev_stage="tailbud embryo"
                     /clone_lib="Nori Satoh unpublished cDNA library, tailbud
                     embryo"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 5;  Length 528;
  Best Local Similarity   75.0%;  Pred. No. 4e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||| || ||:|:|||:
Db        160 TTATAAGTGTGGATGTCCAT 179


RESULT 15
BW392185/c
LOCUS       BW392185                551 bp    mRNA    linear   EST 28-MAY-2004
DEFINITION  BW392185 Yutaka Satou unpublished cDNA library, embryo whole animal
            Ciona intestinalis cDNA clone ciem807o10 3', mRNA sequence.
ACCESSION   BW392185
VERSION     BW392185.1  GI:47808013
KEYWORDS    EST.
SOURCE      Ciona intestinalis
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 551)
  AUTHORS   Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
  TITLE     Expressed genes in Ciona intestinalis (2004)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Yutaka Satou
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4095
            Fax: 81-75-705-1113
            Email: yutaka@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .551
                     /organism="Ciona intestinalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:7719"
                     /clone="ciem807o10"
                     /tissue_type="whole animal"
                     /dev_stage="embryo"
                     /clone_lib="Yutaka Satou unpublished cDNA library, embryo
                     whole animal"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 5;  Length 551;
  Best Local Similarity   75.0%;  Pred. No. 4e+02;
  Matches   15;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TTATAAGGGTCGAUGUCCAU 20
              ||||||| || ||:|:|||:
Db        319 TTATAAGTGTGGATGTCCAT 300


Search completed: March 17, 2005, 11:07:52
Job time : 1389.27 secs
```

THIS PAGE LEFT BLANK

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-65
Perfect score: 20
Sequence: 1 aaattctttataagggucga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	6	AR027821	AR027821 Sequence
2	20	100.0	27	6	AR027819	AR027819 Sequence
c 3	20	100.0	28	6	AR103926	AR103926 Sequence
c 4	20	100.0	35	6	BD236992	BD236992 DNA vacci
c 5	20	100.0	81	6	I92348	I92348 Sequence 9
c 6	20	100.0	253	14	AY329529	AY329529 Hepatitis
c 7	20	100.0	253	14	AY329561	AY329561 Hepatitis
c 8	20	100.0	253	14	AY329562	AY329562 Hepatitis
c 9	20	100.0	253	14	AY329568	AY329568 Hepatitis
c 10	20	100.0	253	14	AY329573	AY329573 Hepatitis
c 11	20	100.0	253	14	AY329575	AY329575 Hepatitis
c 12	20	100.0	253	14	AY329581	AY329581 Hepatitis
c 13	20	100.0	294	14	AF390000	AF390000 Hepatitis
c 14	20	100.0	333	14	HPBHBED	L12359 Hepatitis B
c 15	20	100.0	398	14	AB167603	AB167603 Hepatitis
c 16	20	100.0	398	14	AB167637	AB167637 Hepatitis
c 17	20	100.0	406	14	AB163815	AB163815 Hepatitis
c 18	20	100.0	406	14	AB163817	AB163817 Hepatitis
c 19	20	100.0	439	14	AY254503	AY254503 Hepatitis
c 20	20	100.0	456	14	AY509204	AY509204 Hepatitis
c 21	20	100.0	488	14	AY274419	AY274419 Hepatitis
c 22	20	100.0	488	14	AY274420	AY274420 Hepatitis
c 23	20	100.0	488	14	AY274422	AY274422 Hepatitis
c 24	20	100.0	488	14	AY274424	AY274424 Hepatitis
c 25	20	100.0	488	14	AY274427	AY274427 Hepatitis
c 26	20	100.0	488	14	AY274428	AY274428 Hepatitis
c 27	20	100.0	488	14	AY274429	AY274429 Hepatitis
c 28	20	100.0	488	14	AY274430	AY274430 Hepatitis
c 29	20	100.0	488	14	AY274431	AY274431 Hepatitis
c 30	20	100.0	488	14	AY274432	AY274432 Hepatitis
c 31	20	100.0	488	14	AY274433	AY274433 Hepatitis
c 32	20	100.0	488	14	AY274434	AY274434 Hepatitis
c 33	20	100.0	488	14	AY274436	AY274436 Hepatitis
c 34	20	100.0	548	14	AY382500	AY382500 Hepatitis
c 35	20	100.0	548	14	AY382501	AY382501 Hepatitis
c 36	20	100.0	548	14	AY382502	AY382502 Hepatitis
c 37	20	100.0	548	14	AY382521	AY382521 Hepatitis
c 38	20	100.0	548	14	AY382522	AY382522 Hepatitis
c 39	20	100.0	548	14	AY382523	AY382523 Hepatitis
c 40	20	100.0	548	14	AY382524	AY382524 Hepatitis
c 41	20	100.0	548	14	AY382525	AY382525 Hepatitis
c 42	20	100.0	548	14	AY382526	AY382526 Hepatitis
c 43	20	100.0	548	14	AY382527	AY382527 Hepatitis
c 44	20	100.0	552	6	BD236991	BD236991 DNA vacci
c 45	20	100.0	552	14	AB023678	AB023678 Hepatitis

ALIGNMENTS

RESULT 1
AR027821
LOCUS AR027821 20 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 19 from patent US 5856459.
ACCESSION AR027821
VERSION AR027821.1 GI:5938641
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and Mills,J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 19 05-JAN-1999;
FEATURES Location/Qualifiers
source 1. .20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATTCTTTATAAGGGUCGA 20
||||||||||||||||:|||
Db 1 AAATTCTTTATAAGGGTCGA 20

RESULT 2
AR027819
LOCUS AR027819 27 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 17 from patent US 5856459.
ACCESSION AR027819
VERSION AR027819.1 GI:5938639
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 27)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and

```
            Mills,J.S.
  TITLE     Oligonucleotides specific for hepatitis B virus
  JOURNAL   Patent: US 5856459-A 17 05-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 27;
  Best Local Similarity   95.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db          1 AAATTCTTTATAAGGGTCGA 20


RESULT 3
AR103926/c
LOCUS       AR103926                  28 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 1 from patent US 6087556.
ACCESSION   AR103926
VERSION     AR103926.1  GI:12815514
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Feitelson,M. and Siracusa,L.
  TITLE     Transgenic animals capable of replicating hepatitis viruses and
            mimicking chronic hepatitis infection in humans
  JOURNAL   Patent: US 6087556-A 1 11-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 28;
  Best Local Similarity   95.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db         27 AAATTCTTTATAAGGGTCGA 8


RESULT 4
BD236992/c
LOCUS       BD236992                  35 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  DNA vaccination to cholesterol ester transfer protein in the
            treatment of atherosclerosis.
ACCESSION   BD236992
VERSION     BD236992.1  GI:33046762
KEYWORDS    JP 2002516656-A/17.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Needleman,P. and Glenn,K.
  TITLE     DNA vaccination to cholesterol ester transfer protein in the
            treatment of atherosclerosis
  JOURNAL   Patent: JP 2002516656-A 17 11-JUN-2002;
            MONSANTO CO
COMMENT     OS   Unidentified
            PN   JP 2002516656-A/17
            PD   11-JUN-2002
            PF   17-SEP-1998 JP 2000512947
            PR   19-SEP-1997 US    08/934367
            PI   PHILIP NEEDLEMAN,KEVIN GLENN
            PC   C12N15/09,A61K48/00,C12N15/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   DNA vaccination to cholesterol ester transfer protein in the
            CC        treatment of
            CC   atherosclerosis
            FH   Key             Location/Qualifiers
            FT   source          1. .35
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 35;
  Best Local Similarity   95.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db         33 AAATTCTTTATAAGGGTCGA 14


RESULT 5
I92348/c
LOCUS       I92348                    81 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 9 from patent US 5728518.
ACCESSION   I92348
VERSION     I92348.1  GI:3936818
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Carmichael,E.
  TITLE     Antiviral poly-and oligonucleotides
  JOURNAL   Patent: US 5728518-A 9 17-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .81
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 81;
  Best Local Similarity   95.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db         81 AAATTCTTTATAAGGGTCGA 62


RESULT 6
AY329529/c
LOCUS       AY329529                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate A611252E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329529
VERSION     AY329529.1  GI:37625315
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="A611252E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95861.1"
                     /db_xref="GI:37625316"
                     /translation="STTDLEAYFKDCVFKDWEELGEETRLMIFVLGGCRHKLVCAPAS
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95862.1"
                     /db_xref="GI:37625317"
                     /translation="MQLFHLCLVISCTCPTFQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   95.0%;  Pred. No. 89;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        247 AAATTCTTTATAAGGGTCGA 228


RESULT 7
AY329561/c
LOCUS       AY329561                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D272811E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329561
VERSION     AY329561.1  GI:37625410
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D272811E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95924.1"
                     /db_xref="GI:37625411"
                     /translation="STTDLEAYFKDCLFKDWEELGEETRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95925.1"
                     /db_xref="GI:37625412"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   95.0%;  Pred. No. 89;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        247 AAATTCTTTATAAGGGTCGA 228


RESULT 8
AY329562/c
LOCUS       AY329562                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D273984E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329562
VERSION     AY329562.1  GI:37625413
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D273984E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95926.1"
                     /db_xref="GI:37625414"
                     /translation="STTDLEAYFKDCLFKDWEELGEEIRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95927.1"
                     /db_xref="GI:37625415"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   95.0%;  Pred. No. 89;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        247 AAATTCTTTATAAGGGTCGA 228


RESULT 9
AY329568/c
```

```
LOCUS       AY329568                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D296668E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329568
VERSION     AY329568.1  GI:37625431
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D296668E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95938.1"
                     /db_xref="GI:37625432"
                     /translation="STTDLEAYFKDCLFKDWEELGEELRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95939.1"
                     /db_xref="GI:37625433"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   95.0%;  Pred. No. 89;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        247 AAATTCTTTATAAGGGTCGA 228


RESULT 10
AY329573/c
LOCUS       AY329573                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D604917E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329573
VERSION     AY329573.1  GI:37625446
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D604917E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95948.1"
                     /db_xref="GI:37625447"
                     /translation="STTDLEAYFKDCLFKDWEELGEEIRLKVFVLGGCRHKLVCAPAS
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95949.1"
                     /db_xref="GI:37625448"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   95.0%;  Pred. No. 89;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        247 AAATTCTTTATAAGGGTCGA 228


RESULT 11
AY329575/c
LOCUS       AY329575                253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D611058E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329575
VERSION     AY329575.1  GI:37625452
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D611058E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95952.1"
                     /db_xref="GI:37625453"
                     /translation="STTDLEAYFKDCLFKDWEELGEETRLMIFVLGGCRHKLVCAPAP
                     CNFFTSA"
```

```
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95953.1"
                     /db_xref="GI:37625454"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   95.0%;  Pred. No. 89;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        247 AAATTCTTTATAAGGGTCGA 228


RESULT 12
AY329581/c
LOCUS       AY329581                 253 bp    DNA     linear   VRL 08-JUN-2004
DEFINITION  Hepatitis B virus isolate D639472E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION   AY329581
VERSION     AY329581.1  GI:37625470
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 253)
  AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
  TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
  JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
   PUBMED   15184419
REFERENCE   2  (bases 1 to 253)
  AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES             Location/Qualifiers
     source          1. .253
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D639472E"
                     /db_xref="taxon:10407"
     CDS             <1. .158
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAQ95964.1"
                     /db_xref="GI:37625471"
                     /translation="STTDLEAYFKDCLFKDWEELGEELRLLIFVLGGCRHKLVCAPAP
                     CNFFTSA"
     CDS             134. .217
                     /codon_start=1
                     /product="preC/C protein"
                     /protein_id="AAQ95965.1"
                     /db_xref="GI:37625472"
                     /translation="MQLFHLCLIISCSCPTVQASKLCLGWL"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 253;
  Best Local Similarity   95.0%;  Pred. No. 89;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        247 AAATTCTTTATAAGGGTCGA 228


RESULT 13
AF390000/c
LOCUS       AF390000                 294 bp    DNA     linear   VRL 06-MAR-2002
DEFINITION  Hepatitis B virus isolate D3 X protein gene, partial cds; and
            nonfunctional precore/core protein gene, partial sequence.
ACCESSION   AF390000
VERSION     AF390000.1  GI:16266099
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   Castro,L.D., Niel,C. and Gomes,S.A.
  TITLE     Low frequency of mutations in the core promoter and precore regions
            of hepatitis B virus in anti-HBe positive Brazilian carriers
  JOURNAL   BMC Microbiol. 1 (1), 10 (2001)
   PUBMED   11472634
REFERENCE   2  (bases 1 to 294)
  AUTHORS   De Castro,L., Niel,C. and Gomes,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-2001) Virology, FIOCRUZ, Av. Brasil 4365, Rio de
            Janeiro, RJ 21045-900, Brazil
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="D3"
                     /db_xref="taxon:10407"
     CDS             <1. .119
                     /codon_start=3
                     /product="X protein"
                     /protein_id="AAL16752.1"
                     /db_xref="GI:16266100"
                     /translation="FKDWEELGDDSRLMIYVLGGCRHKLVCAPAPCNFFTSA"
     misc_feature    95. .>294
                     /note="nonfunctional precore/core protein due to mutation"
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 294;
  Best Local Similarity   95.0%;  Pred. No. 87;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        208 AAATTCTTTATAAGGGTCGA 189


RESULT 14
HPBHBED/c
LOCUS       HPBHBED                  333 bp    DNA     linear   VRL 28-JUN-1996
DEFINITION  Hepatitis B virus precore and core protein gene, 5' end of cds.
ACCESSION   L12359
VERSION     L12359.1  GI:306267
KEYWORDS    HBcAg protein; HBeAg protein; core protein; nucleotide binding
            protein; precore protein.
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (sites)
  AUTHORS   Tong,S.P., Li,J.S., Vitvitski,L. and Trepo,C.
  TITLE     Active hepatitis B virus replication in the presence of anti-HBe is
            associated with viral variants containing an inactive pre-C region
  JOURNAL   Virology 176 (2), 596-603 (1990)
  MEDLINE   90266476
   PUBMED   2345966
REFERENCE   2  (bases 1 to 333)
  AUTHORS   Li,J.S., Tong,S.P., Wen,Y.M., Vitvitski,L., Zhang,Q. and Trepo,C.
  TITLE     Hepatitis B virus genotype A rarely circulates as an HBe-minus
            mutant: possible contribution of a single nucleotide in the precore
            region
  JOURNAL   J. Virol. 67 (9), 5402-5410 (1993)
  MEDLINE   93353617
   PUBMED   8350403
COMMENT     Original source text: Hepatitis B virus DNA.
```

```
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
                     /note="genotype D; from French HBeAg- patient"
     CDS             1. .>333
                     /standard_name="HBeAg"
                     /pseudo
                     /codon_start=1
                     /product="precore protein"
     variation       82. .84
                     /note="HBe-abolishing"
     CDS             88. .>333
                     /standard_name="HBcAb"
                     /codon_start=1
                     /protein_id="AAB03102.1"
                     /db_xref="GI:306268"
                     /translation="MDIDPYKEFGATVELLTFLPSDFFPSVRDLLDTAAALYREALES
                     PEHCSHHHTALRQAILCWGELMTLATWVGANLDDPASR"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 333;
  Best Local Similarity   95.0%;  Pred. No. 85;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        114 AAATTCTTTATAAGGGTCGA 95


RESULT 15
AB167603/c
LOCUS       AB167603                398 bp    DNA     linear   VRL 01-OCT-2004
DEFINITION  Hepatitis B virus gene for polyprotein, partial cds, clone: NEP75.
ACCESSION   AB167603
VERSION     AB167603.1  GI:53148166
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1
  AUTHORS   Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
  TITLE     A Case-control Study for Differences among Hepatitis B Virus
            Infections of Genotypes A (Subtypes Aa and Ae) and D
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 398)
  AUTHORS   Tanaka,Y. and Mizokami,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-2004) Yasuhito Tanaka, Nagoya City University
            Graduate School of Medical Sciences, Department of Clinical
            Molecular Informative Medicine; 1 Kawasumi, Mizuho-cho, Mizuho-ku,
            Nagoya, Aichi 467-8601, Japan (E-mail:ytanaka@med.nagoya-cu.ac.jp,
            Tel:81-52-853-8292, Fax:81-52-842-0021)
FEATURES             Location/Qualifiers
     source          1. .398
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10407"
                     /clone="NEP75"
     CDS             184. .>398
                     /note="precore-core region"
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="BAD52175.1"
                     /db_xref="GI:53148167"
                     /translation="MQLFHLCLIISCTCPTFQASKLCLGWLWGMDIDPYKEFGATVEL
                     LSFLPSDFFPSVRDLLDTASALYREAL"
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 398;
  Best Local Similarity   95.0%;  Pred. No. 82;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db        297 AAATTCTTTATAAGGGTCGA 278


Search completed: March 17, 2005, 08:14:18
Job time : 683.733 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-65
Perfect score: 20
Sequence: 1 aaattctttataagggucga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	2	AAT72579	Aat72579 Hepatitis
2	20	100.0	20	2	AAT72580	Aat72580 Hepatitis
3	20	100.0	27	2	AAT72577	Aat72577 Hepatitis
c 4	20	100.0	28	2	AAQ80499	Aaq80499 Primer to
c 5	20	100.0	28	2	AAV45201	Aav45201 Primer MF
c 6	20	100.0	28	3	AAA62585	Aaa62585 Transgeni
c 7	20	100.0	28	12	ADN36074	Adn36074 Probe #15
c 8	20	100.0	28	12	ADN36080	Adn36080 Probe #16
c 9	20	100.0	28	12	ADN36073	Adn36073 Probe #15
c 10	20	100.0	28	12	ADN36075	Adn36075 Probe #15
c 11	20	100.0	28	12	ADN36078	Adn36078 Probe #15
c 12	20	100.0	28	12	ADN36077	Adn36077 Probe #15
c 13	20	100.0	28	12	ADN36076	Adn36076 Probe #15
c 14	20	100.0	28	12	ADN36072	Adn36072 Probe #15
c 15	20	100.0	28	12	ADN36079	Adn36079 Probe #16
c 16	20	100.0	35	2	AAX36590	Aax36590 PCR prime
c 17	20	100.0	35	8	ABX95880	Abx95880 PCR prime
c 18	20	100.0	35	10	ACD07807	Acd07807 Hepatitis
c 19	20	100.0	40	10	ADE10983	Ade10983 Chimeric
c 20	20	100.0	40	10	ADE10985	Ade10985 Chimeric
c 21	20	100.0	40	12	ADG64056	Adg64056 Recombina
c 22	20	100.0	40	12	ADG64054	Adg64054 Recombina
c 23	20	100.0	40	12	ADP73665	Adp73665 HBV pKK-2
c 24	20	100.0	40	12	ADP73667	Adp73667 HBV pKK-2
c 25	20	100.0	40	13	ADR12912	Adr12912 HBV Hbc-C
c 26	20	100.0	40	13	ADR12910	Adr12910 HBV Hbc-C
c 27	20	100.0	41	10	ADG46976	Adg46976 PCR prime
c 28	20	100.0	41	11	ADM83221	Adm83221 PCR prime
c 29	20	100.0	52	10	ADE11060	Ade11060 Chimeric
c 30	20	100.0	52	10	ADG47004	Adg47004 PCR prime
c 31	20	100.0	52	11	ADM83249	Adm83249 Influenza
c 32	20	100.0	52	12	ADG64131	Adg64131 Recombina
c 33	20	100.0	52	12	ADP73784	Adp73784 Influenza
c 34	20	100.0	52	13	ADR12987	Adr12987 Influenza
c 35	20	100.0	53	12	ADN36055	Adn36055 Probe #13
c 36	20	100.0	59	10	ADE11058	Ade11058 Chimeric
c 37	20	100.0	59	10	ADG47002	Adg47002 PCR prime
c 38	20	100.0	59	11	ADM83247	Adm83247 Influenza
c 39	20	100.0	59	12	ADG64129	Adg64129 Recombina
c 40	20	100.0	59	12	ADP73782	Adp73782 Influenza
c 41	20	100.0	59	13	ADR12985	Adr12985 Influenza
42	20	100.0	504	11	ADM41005	Adm41005 HBc relat
c 43	20	100.0	504	11	ADM41004	Adm41004 HBc relat
c 44	20	100.0	513	6	ABK67524	Abk67524 DNA encod
c 45	20	100.0	513	6	ABK67525	Abk67525 DNA encod

ALIGNMENTS

RESULT 1
AAT72579
ID AAT72579 standard; DNA; 20 BP.
XX
AC AAT72579;
XX
DT 04-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV93b.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1. .20
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
XX
PN WO9639502-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP002432.
XX
PR 06-JUN-1995; 95US-00467397.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (HYBR-) HYBRIDON INC.
XX
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX
DR WPI; 1997-043124/04.
XX
PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
PS Claim 1; Page 12; 81pp; English.
XX
CC The present sequence represents a synthetic oligonucleotide HBV93b which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a

```
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 7 A; 2 C; 4 G; 7 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 7;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db          1 AAATTCTTTATAAGGGTCGA 20


RESULT 2
AAT72580
ID   AAT72580 standard; DNA; 20 BP.
XX
AC   AAT72580;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV93Mb.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .20
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
FT   misc_RNA        16. .20
FT                   /*tag=  b
FT                   /note= "2'-OMe RNA"
FT   modified_base   16
FT                   /*tag=  c
FT                   /mod_base= gm
FT   modified_base   17
FT                   /*tag=  d
FT                   /mod_base= um
FT   modified_base   18
FT                   /*tag=  e
FT                   /mod_base= cm
FT   modified_base   19
FT                   /*tag=  f
FT                   /mod_base= gm
FT   modified_base   20
FT                   /*tag=  g
FT                   /mod_base= OTHER
FT                   /note= "2'-O-methyladenosine"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV93Mb which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 20 BP; 7 A; 2 C; 4 G; 6 T; 1 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||||||
Db          1 AAATTCTTTATAAGGGUCGA 20


RESULT 3
AAT72577
ID   AAT72577 standard; DNA; 27 BP.
XX
AC   AAT72577;
XX
DT   04-SEP-1997  (first entry)
XX
DE   Hepatitis B virus RNA antisense oligonucleotide HBV94b.
XX
KW   HBV; HBV infection; inhibition; replication; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .27
FT                   /*tag=  a
FT                   /note= "Internucleotide linkages are phosphorothioate"
XX
PN   WO9639502-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP002432.
XX
PR   06-JUN-1995;   95US-00467397.
XX
PA   (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA   (HYBR-) HYBRIDON INC.
XX
PI   Craig CJ,  Frank BL,  Goodchild J,  Jupp R,  Kilkuskie RE,  Mills JS;
PI   Roberts NA,  Roberts PC,  Slade A;
XX
DR   WPI; 1997-043124/04.
XX
PT   Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT   used in the detection and treatment of HBV infection.
XX
PS   Claim 1; Page 12; 81pp; English.
XX
CC   The present sequence represents a synthetic oligonucleotide HBV94b which
CC   is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC   antisense oligonucleotide may be used to detect the presence of HBV in a
CC   sample. The antisense oligonucleotide, and oligonucleotides containing a
CC   sequence which is complementary to at least two non- contiguous regions
CC   of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC   cell or for the treatment of HBV infection
XX
SQ   Sequence 27 BP; 8 A; 4 C; 5 G; 10 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 27;
  Best Local Similarity   95.0%;  Pred. No. 7.1;
```

```
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db          1 AAATTCTTTATAAGGGTCGA 20


RESULT 4
AAQ80499/c
ID   AAQ80499 standard; DNA; 28 BP.
XX
AC   AAQ80499;
XX
DT   25-MAR-2003  (revised)
DT   23-AUG-1995  (first entry)
XX
DE   Primer to amplify hepatitis B virus core region.
XX
KW   hepatitis B virus; X region; core region; primer; PCR; amplification;
KW   polymerase chain reaction; detection; viral infection; ss.
XX
OS   Synthetic.
XX
PN   WO9429483-A1.
XX
PD   22-DEC-1994.
XX
PF   03-JUN-1994;   94WO-US006360.
XX
PR   08-JUN-1993;   93US-00074346.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Feitelson M,  Duan L,  Guo J;
XX
DR   WPI; 1995-036505/05.
XX
PT   Detection of hepatitis B virus (HBV) variants having deletions in the X
PT   region - by detection of antibodies against HBV polymerase and HB X
PT   antigen.
XX
PS   Claim 3; Page 34; 45pp; English.
XX
CC   This primer designated MF03 covers nucleotide bases 1903-1949 at the
CC   beginning of the hepatitis B virus (HBV) core open reading frame. It is
CC   used with MFO4 (AAQ80500) to amplify the core gene. The primers allow the
CC   detection of a specific class of HBV variants. They are useful for
CC   demonstrating the presence of productive virus infection and may prove
CC   useful in monitoring therapeutics. (Updated on 25-MAR-2003 to correct PN
CC   field.)
XX
SQ   Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   95.0%;  Pred. No. 7.1;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db         27 AAATTCTTTATAAGGGTCGA 8


RESULT 5
AAV45201/c
ID   AAV45201 standard; DNA; 28 BP.
XX
AC   AAV45201;
XX
DT   19-OCT-1998  (first entry)
XX
DE   Primer MFO3.
XX
KW   ss; PCR; primer; amplification; viral infection; bacterial infection;
KW   immune response; hepatitis B virus.
XX
OS   Mus sp.
XX
PN   WO9829121-A1.
XX
PD   09-JUL-1998.
XX
PF   02-JAN-1998;   98WO-US004116.
XX
PR   02-JAN-1997;   97US-0034596P.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Michaels F,  Block T;
XX
DR   WPI; 1998-387782/33.
XX
PT   Modulating immune responses in mammals infected with infectious agent(s)
PT   - e.g. to reduce pathogenicity caused by immune responses in cases where
PT   the infectious agent has limited pathogenicity.
XX
PS   Example 2; Page 37; 55pp; English.
XX
CC   The primers AAV45201 and AAV45202 were used to detect the presence of a
CC   HBV genome which had been microinjected into embryos of SCID mice in an
CC   example to demonstrate modulating an immune response in a mammal infected
CC   with an infectious agent. This comprises transmucosal administration of a
CC   composition comprising an epitope which is located in close proximity to
CC   the immune response. The process may be used in treatment of mammals
CC   which are acutely or chronically infected with infectious agents, such as
CC   viruses or bacteria. It may be used to increase the immune response, or
CC   it may be used to decrease the immune response in cases where the
CC   infectious agent itself exhibits limited pathogenicity but the immune
CC   response to the infectious agent causes more significant pathogenicity.
CC   This can be the case in, e.g. hepatitis B virus (HBV) infection. The
CC   process can modulate undesirable autoimmune responses exhibited by
CC   mammals infected with viral, bacterial and parasitic agents. It can
CC   prevent life-long disabilities which result from these infections
XX
SQ   Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   95.0%;  Pred. No. 7.1;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db         27 AAATTCTTTATAAGGGTCGA 8


RESULT 6
AAA62585/c
ID   AAA62585 standard; DNA; 28 BP.
XX
AC   AAA62585;
XX
DT   22-NOV-2000  (first entry)
XX
DE   Transgenic SCID mouse hepatitis virus transgene PCR primer MF03.
XX
KW   Mouse; SCID; severe combined immunodeficiency; transgenic mouse;
KW   hepatitis virus; hepatitis B; hepatitis C; chronic liver disease;
KW   PCR primer; ss.
XX
OS   Mus sp.
XX
PN   US6087556-A.
XX
PD   11-JUL-2000.
XX
PF   07-JAN-1998;   98US-00003200.
XX
```

PR 02-MAY-1996; 96US-00641803.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Siracusa L, Peitelson M;
XX
DR WPI; 2000-523731/47.
XX
PT Transgenic mouse useful in methods for evaluating interactions of
PT chemical, drug or immunomodulating agent with hepatitis virus, lacks
PT functional T-cells and B-cells and is capable of replicating hepatitis
PT viruses.
XX
PS Example 3; Col 8; 7pp; English.
XX
CC The present sequence is a PCR primer used to detect a transgene in severe
CC combined immunodeficient (SCID) mice. Transgenic immunodeficient mice
CC were produced that are not tolerant to hepatitis viral antigens, lack
CC functional T-cells and B-cells and contain integrated hepatitis virus DNA
CC in the somatic and germ cells. The hepatitis virus gene is expressed and
CC the hepatitis virus is replicated in the transgenic mouse. The mouse may
CC be used as an animal model for evaluating interactions of a chemical,
CC drug or immunomodulating agent with a hepatitis virus. It is also useful
CC for the assessment of anti-viral and immunomodulatory intervention
CC therapies, including the screening of drug candidates. It can be used to
CC analyse the virus-host relationship, to evaluate the relationship between
CC the virus and chemicals metabolised and/or detoxified by the liver, and
CC to identify cellular biochemical pathways contributing to the development
CC and progression of chronic liver disease. The transgenic mouse is thus
CC useful for elucidating the effects of hepatitis virus on hepatic
CC metabolism
XX
SQ Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 28;
Best Local Similarity 95.0%; Pred. No. 7.1;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATTCTTTATAAGGGUCGA 20
||||||||||||||||:|||
Db 27 AAATTCTTTATAAGGGTCGA 8

RESULT 7
ADN36074/c
ID ADN36074 standard; DNA; 28 BP.
XX
AC ADN36074;
XX
DT 01-JUL-2004 (first entry)
XX
DE Probe #155 to determine effect of long term lamivudine treatment of HBV.
XX
KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS Hepatitis B virus.
XX
PN WO2004031729-A2.
XX
PD 15-APR-2004.
XX
PF 01-OCT-2003; 2003WO-US031121.
XX
PR 01-OCT-2002; 2002US-0415301P.
XX
PA (GEOU) UNIV GEORGETOWN.
XX
PI Korba BE, Ciancio A, Gerin JL;
XX
DR WPI; 2004-348004/32.
XX
PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS Claim 31; SEQ ID NO 155; 107pp; English.
XX
CC The invention relates to a method of predicting the long term response of
CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC leucine at amino acid position (aa) 91 in the DNA polymerase region
CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC 604) in the DNA polymerase region of HBV. The method comprises
CC determining whether the HBV carried by the host bears one or more of the
CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC and C1962A representing specific double point mutations in the DNA
CC precore/core promoter or ORF region. The method and kit is useful in
CC predicting the long-term response of a host of HBV to 3TC therapy (also
CC known as lamivudine). This sequence represents an oligonucleotide
CC sequence used in the method of the invention to detect a mutation in the
CC above mentioned sequences.
XX
SQ Sequence 28 BP; 9 A; 5 C; 6 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 95.0%; Pred. No. 7.1;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATTCTTTATAAGGGUCGA 20
||||||||||||||||:|||
Db 26 AAATTCTTTATAAGGGTCGA 7

RESULT 8
ADN36080/c
ID ADN36080 standard; DNA; 28 BP.
XX
AC ADN36080;
XX
DT 01-JUL-2004 (first entry)
XX
DE Probe #161 to determine effect of long term lamivudine treatment of HBV.
XX
KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS Hepatitis B virus.
XX
PN WO2004031729-A2.
XX
PD 15-APR-2004.
XX
PF 01-OCT-2003; 2003WO-US031121.
XX
PR 01-OCT-2002; 2002US-0415301P.
XX
PA (GEOU) UNIV GEORGETOWN.
XX
PI Korba BE, Ciancio A, Gerin JL;
XX
DR WPI; 2004-348004/32.
XX
PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS Claim 31; SEQ ID NO 161; 107pp; English.
XX
CC The invention relates to a method of predicting the long term response of
CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC leucine at amino acid position (aa) 91 in the DNA polymerase region
CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon

CC 604) in the DNA polymerase region of HBV. The method comprises
CC determining whether the HBV carried by the host bears one or more of the
CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC and C1962A representing specific double point mutations in the DNA
CC precore/core promoter or ORF region. The method and kit is useful in
CC predicting the long-term response of a host of HBV to 3TC therapy (also
CC known as lamivudine). This sequence represents an oligonucleotide
CC sequence used in the method of the invention to detect a mutation in the
CC above mentioned sequences.
XX
SQ Sequence 28 BP; 8 A; 6 C; 5 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 95.0%; Pred. No. 7.1;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATTCTTTATAAGGGUCGA 20
||||||||||||||||:|||
Db 20 AAATTCTTTATAAGGGTCGA 1

RESULT 9
ADN36073/c
ID ADN36073 standard; DNA; 28 BP.
XX
AC ADN36073;
XX
DT 01-JUL-2004 (first entry)
XX
DE Probe #154 to determine effect of long term lamivudine treatment of HBV.
XX
KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS Hepatitis B virus.
XX
PN WO2004031729-A2.
XX
PD 15-APR-2004.
XX
PF 01-OCT-2003; 2003WO-US031121.
XX
PR 01-OCT-2002; 2002US-0415301P.
XX
PA (GEOU) UNIV GEORGETOWN.
XX
PI Korba BE, Ciancio A, Gerin JL;
XX
DR WPI; 2004-348004/32.
XX
PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS Claim 31; SEQ ID NO 154; 107pp; English.
XX
CC The invention relates to a method of predicting the long term response of
CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC leucine at amino acid position (aa) 91 in the DNA polymerase region
CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC 604) in the DNA polymerase region of HBV. The method comprises
CC determining whether the HBV carried by the host bears one or more of the
CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC and C1962A representing specific double point mutations in the DNA
CC precore/core promoter or ORF region. The method and kit is useful in
CC predicting the long-term response of a host of HBV to 3TC therapy (also
CC known as lamivudine). This sequence represents an oligonucleotide
CC sequence used in the method of the invention to detect a mutation in the
CC above mentioned sequences.
XX
SQ Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 95.0%; Pred. No. 7.1;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATTCTTTATAAGGGUCGA 20
||||||||||||||||:|||
Db 27 AAATTCTTTATAAGGGTCGA 8

RESULT 10
ADN36075/c
ID ADN36075 standard; DNA; 28 BP.
XX
AC ADN36075;
XX
DT 01-JUL-2004 (first entry)
XX
DE Probe #156 to determine effect of long term lamivudine treatment of HBV.
XX
KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS Hepatitis B virus.
XX
PN WO2004031729-A2.
XX
PD 15-APR-2004.
XX
PF 01-OCT-2003; 2003WO-US031121.
XX
PR 01-OCT-2002; 2002US-0415301P.
XX
PA (GEOU) UNIV GEORGETOWN.
XX
PI Korba BE, Ciancio A, Gerin JL;
XX
DR WPI; 2004-348004/32.
XX
PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS Claim 31; SEQ ID NO 156; 107pp; English.
XX
CC The invention relates to a method of predicting the long term response of
CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC leucine at amino acid position (aa) 91 in the DNA polymerase region
CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC 604) in the DNA polymerase region of HBV. The method comprises
CC determining whether the HBV carried by the host bears one or more of the
CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC and C1962A representing specific double point mutations in the DNA
CC precore/core promoter or ORF region. The method and kit is useful in
CC predicting the long-term response of a host of HBV to 3TC therapy (also
CC known as lamivudine). This sequence represents an oligonucleotide
CC sequence used in the method of the invention to detect a mutation in the
CC above mentioned sequences.
XX
SQ Sequence 28 BP; 10 A; 5 C; 6 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 95.0%; Pred. No. 7.1;

```
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db         25 AAATTCTTTATAAGGGTCGA 6


RESULT 11
ADN36078/c
ID   ADN36078 standard; DNA; 28 BP.
XX
AC   ADN36078;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Probe #159 to determine effect of long term lamivudine treatment of HBV.
XX
KW   ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS   Hepatitis B virus.
XX
PN   WO2004031729-A2.
XX
PD   15-APR-2004.
XX
PF   01-OCT-2003; 2003WO-US031121.
XX
PR   01-OCT-2002; 2002US-0415301P.
XX
PA   (GEOU ) UNIV GEORGETOWN.
XX
PI   Korba BE,  Ciancio A,  Gerin JL;
XX
DR   WPI; 2004-348004/32.
XX
PT   Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT   3TC therapy comprises determining whether the HBV bears a nucleic acid
PT   encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS   Claim 31; SEQ ID NO 159; 107pp; English.
XX
CC   The invention relates to a method of predicting the long term response of
CC   a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC   the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC   leucine at amino acid position (aa) 91 in the DNA polymerase region
CC   (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC   604) in the DNA polymerase region of HBV. The method comprises
CC   determining whether the HBV carried by the host bears one or more of the
CC   following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC   codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC   T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC   promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC   changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC   and C1962A representing specific double point mutations in the DNA
CC   precore/core promoter or ORF region. The method and kit is useful in
CC   predicting the long-term response of a host of HBV to 3TC therapy (also
CC   known as lamivudine). This sequence represents an oligonucleotide
CC   sequence used in the method of the invention to detect a mutation in the
CC   above mentioned sequences.
XX
SQ   Sequence 28 BP; 9 A; 6 C; 5 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 12;  Length 28;
  Best Local Similarity   95.0%;  Pred. No. 7.1;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db         22 AAATTCTTTATAAGGGTCGA 3


RESULT 12
ADN36077/c
ID   ADN36077 standard; DNA; 28 BP.
XX
AC   ADN36077;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Probe #158 to determine effect of long term lamivudine treatment of HBV.
XX
KW   ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS   Hepatitis B virus.
XX
PN   WO2004031729-A2.
XX
PD   15-APR-2004.
XX
PF   01-OCT-2003; 2003WO-US031121.
XX
PR   01-OCT-2002; 2002US-0415301P.
XX
PA   (GEOU ) UNIV GEORGETOWN.
XX
PI   Korba BE,  Ciancio A,  Gerin JL;
XX
DR   WPI; 2004-348004/32.
XX
PT   Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT   3TC therapy comprises determining whether the HBV bears a nucleic acid
PT   encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS   Claim 31; SEQ ID NO 158; 107pp; English.
XX
CC   The invention relates to a method of predicting the long term response of
CC   a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC   the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC   leucine at amino acid position (aa) 91 in the DNA polymerase region
CC   (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC   604) in the DNA polymerase region of HBV. The method comprises
CC   determining whether the HBV carried by the host bears one or more of the
CC   following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC   codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC   T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC   promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC   changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC   and C1962A representing specific double point mutations in the DNA
CC   precore/core promoter or ORF region. The method and kit is useful in
CC   predicting the long-term response of a host of HBV to 3TC therapy (also
CC   known as lamivudine). This sequence represents an oligonucleotide
CC   sequence used in the method of the invention to detect a mutation in the
CC   above mentioned sequences.
XX
SQ   Sequence 28 BP; 10 A; 6 C; 5 G; 7 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 20;  DB 12;  Length 28;
  Best Local Similarity   95.0%;  Pred. No. 7.1;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db         23 AAATTCTTTATAAGGGTCGA 4


RESULT 13
ADN36076/c
ID   ADN36076 standard; DNA; 28 BP.
XX
AC   ADN36076;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Probe #157 to determine effect of long term lamivudine treatment of HBV.
XX
KW   ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
```

XX
OS Hepatitis B virus.
XX
PN WO2004031729-A2.
XX
PD 15-APR-2004.
XX
PF 01-OCT-2003; 2003WO-US031121.
XX
PR 01-OCT-2002; 2002US-0415301P.
XX
PA (GEOU) UNIV GEORGETOWN.
XX
PI Korba BE, Ciancio A, Gerin JL;
XX
DR WPI; 2004-348004/32.
XX
PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS Claim 31; SEQ ID NO 157; 107pp; English.
XX
CC The invention relates to a method of predicting the long term response of
CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC leucine at amino acid position (aa) 91 in the DNA polymerase region
CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC 604) in the DNA polymerase region of HBV. The method comprises
CC determining whether the HBV carried by the host bears one or more of the
CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC and C1962A representing specific double point mutations in the DNA
CC precore/core promoter or ORF region. The method and kit is useful in
CC predicting the long-term response of a host of HBV to 3TC therapy (also
CC known as lamivudine). This sequence represents an oligonucleotide
CC sequence used in the method of the invention to detect a mutation in the
CC above mentioned sequences.
XX
SQ Sequence 28 BP; 10 A; 5 C; 6 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 95.0%; Pred. No. 7.1;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATTCTTTATAAGGGUCGA 20
||||||||||||||||:|||
Db 24 AAATTCTTTATAAGGGTCGA 5

RESULT 14
ADN36072/c
ID ADN36072 standard; DNA; 28 BP.
XX
AC ADN36072;
XX
DT 01-JUL-2004 (first entry)
XX
DE Probe #153 to determine effect of long term lamivudine treatment of HBV.
XX
KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS Hepatitis B virus.
XX
PN WO2004031729-A2.
XX
PD 15-APR-2004.
XX
PF 01-OCT-2003; 2003WO-US031121.
XX
PR 01-OCT-2002; 2002US-0415301P.
XX
PA (GEOU) UNIV GEORGETOWN.
XX
PI Korba BE, Ciancio A, Gerin JL;
XX
DR WPI; 2004-348004/32.
XX
PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS Claim 31; SEQ ID NO 153; 107pp; English.
XX
CC The invention relates to a method of predicting the long term response of
CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC leucine at amino acid position (aa) 91 in the DNA polymerase region
CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC 604) in the DNA polymerase region of HBV. The method comprises
CC determining whether the HBV carried by the host bears one or more of the
CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC and C1962A representing specific double point mutations in the DNA
CC precore/core promoter or ORF region. The method and kit is useful in
CC predicting the long-term response of a host of HBV to 3TC therapy (also
CC known as lamivudine). This sequence represents an oligonucleotide
CC sequence used in the method of the invention to detect a mutation in the
CC above mentioned sequences.
XX
SQ Sequence 28 BP; 10 A; 6 C; 4 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 95.0%; Pred. No. 7.1;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATTCTTTATAAGGGUCGA 20
||||||||||||||||:|||
Db 28 AAATTCTTTATAAGGGTCGA 9

RESULT 15
ADN36079/c
ID ADN36079 standard; DNA; 28 BP.
XX
AC ADN36079;
XX
DT 01-JUL-2004 (first entry)
XX
DE Probe #160 to determine effect of long term lamivudine treatment of HBV.
XX
KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX
OS Hepatitis B virus.
XX
PN WO2004031729-A2.
XX
PD 15-APR-2004.
XX
PF 01-OCT-2003; 2003WO-US031121.
XX
PR 01-OCT-2002; 2002US-0415301P.
XX
PA (GEOU) UNIV GEORGETOWN.
XX
PI Korba BE, Ciancio A, Gerin JL;
XX
DR WPI; 2004-348004/32.
XX
PT Predicting the long-term response of a host of hepatitis B virus (HBV) to

PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
PS Claim 31; SEQ ID NO 160; 107pp; English.
XX
CC The invention relates to a method of predicting the long term response of
CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
CC leucine at amino acid position (aa) 91 in the DNA polymerase region
CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
CC 604) in the DNA polymerase region of HBV. The method comprises
CC determining whether the HBV carried by the host bears one or more of the
CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T196IA
CC and C1962A representing specific double point mutations in the DNA
CC precore/core promoter or ORF region. The method and kit is useful in
CC predicting the long-term response of a host of HBV to 3TC therapy (also
CC known as lamivudine). This sequence represents an oligonucleotide
CC sequence used in the method of the invention to detect a mutation in the
CC above mentioned sequences.
XX
SQ Sequence 28 BP; 9 A; 5 C; 5 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 95.0%; Pred. No. 7.1;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||||||||||||||:|||
Db         21 AAATTCTTTATAAGGGTCGA 2
```

Search completed: March 17, 2005, 06:48:45
Job time : 172.333 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-65
Perfect score: 20
Sequence: 1 aaattctttataagggucga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	18	90.0	866	9	CL798479	CL798479 OR_CBa000
c	2	17.4	87.0	624	6	CA029935	CA029935 HX05J23r
c	3	17.4	87.0	893	7	CK125419	CK125419 BES182410
	4	17.4	87.0	975	9	CNS06GLD	AL397847 T3 end of
	5	17.4	87.0	1003	9	CNS06J82	AL401256 T3 end of
c	6	17.4	87.0	1028	9	CNS06J81	AL401255 T7 end of
c	7	16.8	84.0	85	9	AG265169	AG265169 Lotus cor
c	8	16.8	84.0	293	2	BB009906	BB009906 BB009906
	9	16.8	84.0	299	1	AV258784	AV258784 AV258784
	10	16.8	84.0	407	6	CA856765	CA856765 PfESToac4
	11	16.8	84.0	462	4	BI926889	BI926889 EST546778
c	12	16.8	84.0	494	5	BP527047	BP527047 BP527047
	13	16.8	84.0	508	9	CE435953	CE435953 tigr-gss-
	14	16.8	84.0	572	9	CR154100	CR154100 Forward s
	15	16.8	84.0	588	6	CD865385	CD865385 AZO2.073N
c	16	16.8	84.0	593	7	CF424535	CF424535 lad32b05.
	17	16.8	84.0	620	7	CK537618	CK537618 rswgb0_00
c	18	16.8	84.0	620	9	AG016570	AG016570 Homo sapi
	19	16.8	84.0	626	3	AY068711	AY068711 Schmidtea
c	20	16.8	84.0	630	9	AG016569	AG016569 Homo sapi
c	21	16.8	84.0	678	9	AG119858	AG119858 Pan trogl
c	22	16.8	84.0	700	7	CK569388	CK569388 HO13B17w
c	23	16.8	84.0	707	5	BU041523	BU041523 PP_LEa000
c	24	16.8	84.0	707	9	CE761259	CE761259 tigr-gss-
	25	16.8	84.0	750	9	AG597534	AG597534 Mus muscu
	26	16.8	84.0	752	9	CL748062	CL748062 OR_BBa011
c	27	16.8	84.0	810	9	CC895730	CC895730 ZMMBBb022
	28	16.8	84.0	836	9	CL620368	CL620368 OR_BBa001
	29	16.8	84.0	942	9	CL118472	CL118472 ISB1-70N2
	30	16.8	84.0	1001	9	CNS06N7K	AL406422 T3 end of
c	31	16.8	84.0	1131	8	CC240216	CC240216 CH261-121
	32	16.4	82.0	168	8	BH783463	BH783463 fzmb013f0
c	33	16.4	82.0	312	9	CL834207	CL834207 OR_CBa005
c	34	16.4	82.0	323	9	CL728817	CL728817 OR_BBa006
c	35	16.4	82.0	338	9	CL782534	CL782534 OR_BBa009
c	36	16.4	82.0	350	1	AI556720	AI556720 UI-R-C2p-
c	37	16.4	82.0	431	9	CL611235	CL611235 OR_BBa000
	38	16.4	82.0	467	4	BJ267319	BJ267319 BJ267319
c	39	16.4	82.0	474	9	CL730394	CL730394 OR_BBa006
	40	16.4	82.0	554	9	CL551502	CL551502 OB_ Ba009
	41	16.4	82.0	588	4	BJ271762	BJ271762 BJ271762
	42	16.4	82.0	612	5	BW030883	BW030883 BW030883
	43	16.4	82.0	621	2	BB646609	BB646609 BB646609
c	44	16.4	82.0	637	4	BJ686765	BJ686765 BJ686765
c	45	16.4	82.0	641	4	BJ702377	BJ702377 BJ702377

ALIGNMENTS

RESULT 1
CL798479/c
LOCUS CL798479 866 bp DNA linear GSS 06-AUG-2004
DEFINITION OR_CBa0009F12.f OR_CBa Oryza rufipogon genomic clone OR_CBa0009F12 5', genomic survey sequence.
ACCESSION CL798479
VERSION CL798479.1 GI:51021406
KEYWORDS GSS.
SOURCE Oryza rufipogon
ORGANISM Oryza rufipogon
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 866)
AUTHORS Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0009 row: F column: 12
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .866
/organism="Oryza rufipogon"
/mol_type="genomic DNA"
/db_xref="taxon:4529"
/clone="OR_CBa0009F12"
/tissue_type="young leaves"
/dev_stage="2 week old seedlings"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OR_CBa"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII; drk treated 36 hrs before harvest"
ORIGIN

Query Match 90.0%; Score 18; DB 9; Length 866;

```
  Best Local Similarity   94.4%;  Pred. No. 4.5e+02;
  Matches   17;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUC 18
              ||||||||||||||||:|
Db        811 AAATTCTTTATAAGGGTC 794


RESULT 2
CA029935/c
LOCUS       CA029935                624 bp    mRNA    linear   EST 24-OCT-2002
DEFINITION  HX05J23r HX Hordeum vulgare subsp. vulgare cDNA clone HX05J23
            5-PRIME, mRNA sequence.
ACCESSION   CA029935
VERSION     CA029935.1  GI:24325281
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Zhang,H., Weschke,W., Michalek,W., Stein,N. and Graner,A.
  TITLE     EST sequencing and analysis in barley (2002)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Stein Nils
            Molecular Markers Group, Department Genbank
            Institute of Plant Genetics and Crop Plant Research (IPK)
            Correnesstr. 3, 06466, Gatersleben, Germany
            Tel: 039482-5522
            Fax: 039482-5595
            Email: stein@ipk-gatersleben.de
            Insert Length: 624   Std Error: 0.00
            Plate: 5  row: J  column: 23
            Seq primer: M13rev.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="barke"
                     /sub_species="vulgare"
                     /db_xref="GABI:270606"
                     /db_xref="taxon:112509"
                     /clone="HX05J23"
                     /tissue_type="apex"
                     /dev_stage="apex (3-5 mm in size)"
                     /lab_host="XL10-Gold"
                     /clone_lib="HX"
                     /note="Vector: pBluescript SK+; Site_1: EcoRI (5'-end of
                     cDNA); Site_2: XhoI (3'-end of cDNA); Due to a cloning
                     artefact caused by the kit, in most cases the EcoRI site
                     is NOT present, as well as the EcoRIadapter used for
                     cloning. To excise the insert, restriction sites upstream
                     EcoRI should be used (e.g. BamHI, SalI,PstI). NOTE: Also
                     due to the cloning system used Blue/white selection for
                     recombinats is not 100% reliable."
ORIGIN

  Query Match           87.0%;  Score 17.4;  DB 6;  Length 624;
  Best Local Similarity   89.5%;  Pred. No. 8.5e+02;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCG 19
              ||||||| |||||||:||
Db        473 AAATTCTTCATAAGGGTCG 455


RESULT 3
CK125419/c
LOCUS       CK125419                893 bp    mRNA    linear   EST 01-MAR-2004
DEFINITION  BES1824107p07 BES1824 Hordeum vulgare subsp. vulgare cDNA clone
            MPMGp2010P077 5-PRIME, mRNA sequence.
ACCESSION   CK125419
VERSION     CK125419.1  GI:44808421
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 893)
  AUTHORS   Kramer,A., Feilner,T., Possling,A., Radchuk,V., Weschke,W.,
            Buerkle,L. and Kersten,B.
  TITLE     Application of the protein microarray technology for the
            identification of expression library derived target proteins for
            barley protein kinase CK2
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birgit Kersten* and Winfriede Weschke**
            *Plant Protein Chip Group, Department Lehrach, **Department
            Molecular Genetics, Gene Expression Group
            *Max-Planck-Institute for Molecular Genetics, **Institute of Plant
            Genetics and Crop Plant Research Gatersleben
            *Ihnestr. 73 , D-14195 Berlin, Germany, **Correnesstrasse 3, D-06466
            Gatersleben, Germany
            Tel: *+49(0)30/84131648,**+49(0)394825500
            Fax: *+49(0)30/84131128,**+49(0)394825237
            Email: *kersten@molgen.mpg.de, **weschke@ipk-gatersleben.de
            Insert Length: 893   Std Error: 0.00
            Plate: 7  row: P  column: 7
            Seq primer: pQE65.
FEATURES             Location/Qualifiers
     source          1. .893
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="barke"
                     /sub_species="vulgare"
                     /db_xref="GABI:944977"
                     /db_xref="taxon:112509"
                     /clone="MPMGp2010P077"
                     /tissue_type="embryosac"
                     /dev_stage="0-10 DAF (days after flowering)"
                     /lab_host="E. coli, SCS-1/pSE111"
                     /clone_lib="BES1824"
                     /note="Vector: pQE30NST (AF074376); Site_1: SalI; Site_2:
                     NotI; 0-10 DAF (days after flowering), cDNA synthesis
                     using pBluescript II XR cDNA-library construction kit
                     (Stratagen) with an oligo(dT)-primer containing NotI
                     restriction site and a SalI adapter (Invitrogen). The main
                     library of 21500 clones was rearrayed into the sublibrary
                     BES 1824 containing 4100 putative expression clones. Note:
                     Due to a cloning artefact caused by the kit, in most cases
                     the SalI site is NOT present, as well as the SalI Adapter
                     used for cloning. To excise the insert, restriction sites
                     upstream SalI should be used (e.g. BamHI). Average insert
                     size is 1 kb. Library generation and sequencing was
                     granted in context of GABI; data are also accessible at
                     https://gabi.rzpd.de"
ORIGIN

  Query Match           87.0%;  Score 17.4;  DB 7;  Length 893;
  Best Local Similarity   89.5%;  Pred. No. 8.8e+02;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCG 19
              ||||||| |||||||:||
Db        564 AAATTCTTCATAAGGGTCG 546


RESULT 4
CNS06GLD
LOCUS       CNS06GLD                975 bp    DNA     linear   GSS 30-NOV-2001
DEFINITION  T3 end of clone AS0AA005E07 of library AS0AA from strain CLIB 533
            of Saccharomyces bayanus, genomic survey sequence.
ACCESSION   AL397847
VERSION     AL397847.1  GI:12150829
KEYWORDS    GSS.
```

```
SOURCE      Saccharomyces bayanus
  ORGANISM  Saccharomyces bayanus
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 975)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 975)
  AUTHORS   Bon,E., Neuveglise,C., Casaregola,S., Artiguenave,F., Wincker,P.,
            Aigle,M. and Durrens,P.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 5.
            Saccharomyces bayanus var. uvarum
  JOURNAL   FEBS Lett. 487 (1), 37-41 (2000)
  MEDLINE   20584715
   PUBMED   11152880
REFERENCE   3  (bases 1 to 975)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="Saccharomyces bayanus"
                     /mol_type="genomic DNA"
                     /strain="CLIB 533"
                     /variety="uvarum"
                     /db_xref="taxon:4931"
                     /clone="AS0AA005E07"
                     /clone_lib="AS0AA"
                     /note="end : T3"
     misc_feature    complement(<26. .>958)
                     /note="similar to Saccharomyces cerevisiae ORF YHR172w [
                     SPC97 ; spindle pole body component ]"
                     /evidence=not_experimental
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 975;
  Best Local Similarity   89.5%;  Pred. No. 8.9e+02;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCG 19
              |||||||| |||||||:||
Db        167 AAATTCTTGATAAGGGTCG 185


RESULT 5
CNS06J82
LOCUS       CNS06J82               1003 bp    DNA     linear   GSS 30-NOV-2001
DEFINITION  T3 end of clone AS0AA026F08 of library AS0AA from strain CLIB 533
            of Saccharomyces bayanus, genomic survey sequence.
ACCESSION   AL401256
VERSION     AL401256.1  GI:12158668
KEYWORDS    GSS.
```

```
SOURCE      Saccharomyces bayanus
  ORGANISM  Saccharomyces bayanus
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 1003)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 1003)
  AUTHORS   Bon,E., Neuveglise,C., Casaregola,S., Artiguenave,F., Wincker,P.,
            Aigle,M. and Durrens,P.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 5.
            Saccharomyces bayanus var. uvarum
  JOURNAL   FEBS Lett. 487 (1), 37-41 (2000)
  MEDLINE   20584715
   PUBMED   11152880
REFERENCE   3  (bases 1 to 1003)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .1003
                     /organism="Saccharomyces bayanus"
                     /mol_type="genomic DNA"
                     /strain="CLIB 533"
                     /variety="uvarum"
                     /db_xref="taxon:4931"
                     /clone="AS0AA026F08"
                     /clone_lib="AS0AA"
                     /note="end : T3"
     misc_feature    complement(<3. .>988)
                     /note="similar to Saccharomyces cerevisiae ORF YHR172w [
                     SPC97 ; spindle pole body component ]
                     1 putative frameshift(s)"
                     /evidence=not_experimental
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 1003;
  Best Local Similarity   89.5%;  Pred. No. 8.9e+02;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCG 19
              |||||||| |||||||:||
Db        909 AAATTCTTGATAAGGGTCG 927


RESULT 6
CNS06J81/c
LOCUS       CNS06J81               1028 bp    DNA     linear   GSS 30-NOV-2001
DEFINITION  T7 end of clone AS0AA026F08 of library AS0AA from strain CLIB 533
            of Saccharomyces bayanus, genomic survey sequence.
ACCESSION   AL401255
VERSION     AL401255.1  GI:12158665
```

```
KEYWORDS    GSS.
SOURCE      Saccharomyces bayanus
  ORGANISM  Saccharomyces bayanus
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 1028)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 1028)
  AUTHORS   Bon,E., Neuveglise,C., Casaregola,S., Artiguenave,F., Wincker,P.,
            Aigle,M. and Durrens,P.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 5.
            Saccharomyces bayanus var. uvarum
  JOURNAL   FEBS Lett. 487 (1), 37-41 (2000)
  MEDLINE   20584715
   PUBMED   11152880
REFERENCE   3  (bases 1 to 1028)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .1028
                     /organism="Saccharomyces bayanus"
                     /mol_type="genomic DNA"
                     /strain="CLIB 533"
                     /variety="uvarum"
                     /db_xref="taxon:4931"
                     /clone="AS0AA026F08"
                     /clone_lib="AS0AA"
                     /note="end : T7"
     misc_feature    <8. .>1027
                     /note="similar to Saccharomyces cerevisiae ORF YHR172w [
                     SPC97 ; spindle pole body component ]
                     2 putative frameshift(s)"
                     /evidence=not_experimental
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 1028;
  Best Local Similarity   89.5%;  Pred. No. 8.9e+02;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCG 19
              |||||||| |||||||:||
Db        820 AAATTCTTGATAAGGGTCG 802


RESULT 7
AG265169/c
LOCUS       AG265169                 85 bp    DNA     linear   GSS 22-JUL-2003
DEFINITION  Lotus corniculatus var. japonicus DNA, clone:LjT62e07_sfi, genomic
            survey sequence.
ACCESSION   AG265169
VERSION     AG265169.1  GI:26665008
KEYWORDS    GSS.
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
  ORGANISM  Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Sato,S., Nakamura,Y. and Tabata,S.
  TITLE     Lotus japonicus TAC End sequences
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 85)
  AUTHORS   Sato,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Shusei Sato, Kazusa DNA Research Institute,
            The First Laboratory for Plant Gene Research; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:ssato@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/,
            Tel:81-438-52-3935(ex.2336), Fax:81-438-52-3934)
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Lotus corniculatus var. japonicus"
                     /mol_type="genomic DNA"
                     /strain="Miyakojima MG-20"
                     /variety="japonicus"
                     /db_xref="taxon:34305"
                     /clone="LjT62e07_sfi"
                     /clone_lib="genomic TAC library"
                     /note="VECTOR:pYLTAC7~synonym: Lotus japonicus"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 85;
  Best Local Similarity   85.0%;  Pred. No. 1.4e+03;
  Matches   17;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              |||||||| ||| |||:|||
Db         36 AAATTCTTAATATGGGTCGA 17


RESULT 8
BB009906/c
LOCUS       BB009906                 293 bp    mRNA    linear   EST 22-JUN-2000
DEFINITION  BB009906 RIKEN full-length enriched, 10 day neonate skin Mus
            musculus cDNA clone 4732491N16 3', mRNA sequence.
ACCESSION   BB009906
VERSION     BB009906.1  GI:8130263
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 293)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,
            Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
            Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
            Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
            Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,
            Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A.,
            Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,
            Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
```

```
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
            Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
            Okazaki,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .293
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4732491N16"
                     /sex="mixed"
                     /tissue_type="skin"
                     /dev_stage="10 days neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 10 day neonate
                     skin"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 100.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cloned into the XhoI and BamHI sites.
                     Vector: a modified pBluescript KS(+) after bulk excision
                     from Lambda FLC I"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 2;  Length 293;
  Best Local Similarity   85.0%;  Pred. No. 1.6e+03;
  Matches   17;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              |||||||||||||| |:| |
Db         59 AAATTCTTTATAAGAGTCTA 40


RESULT 9
AV258784
LOCUS       AV258784                299 bp    mRNA    linear   EST 04-NOV-1999
DEFINITION  AV258784 RIKEN full-length enriched, adult male testis (DH10B) Mus
            musculus cDNA clone 4930401P07 3' similar to M18654 Mouse
            testis-specific phosphoglycerate kinase (pgk-2) mRNA, mRNA
            sequence.
ACCESSION   AV258784
VERSION     AV258784.1  GI:6246243
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 299)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
            Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
            Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,
            Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
            Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
            Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K.,
            Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y.,
            Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N.,
            Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A.,
            Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al. 1999)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
            Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and
            Hayashizaki,Y.
             Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
            Okazaki,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4930401P07"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, adult male testis
                     (DH10B)"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cloned into the XhoI and BamHI sites. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 1;  Length 299;
  Best Local Similarity   85.0%;  Pred. No. 1.6e+03;
  Matches   17;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
```

```
              ||||||||||||| |||: ||
Db          6 AAATTCTTTATAGGGGTAGA 25


RESULT 10
CA856765
LOCUS       CA856765                407 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  PfESToac41h07.y1 Plasmodium falciparum 3D7 gametocyte cDNA library
            Plasmodium falciparum 3D7 cDNA 5', mRNA sequence.
ACCESSION   CA856765
VERSION     CA856765.1  GI:27159521
KEYWORDS    EST.
SOURCE      Plasmodium falciparum 3D7
  ORGANISM  Plasmodium falciparum 3D7
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 407)
  AUTHORS   Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D.,
            Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I.,
            Tsagareishvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A.,
            Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D.,
            Waterston,R., Wilson,R. and Sibley,D.
  TITLE     WashU Plasmodium EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: L. David Sibley
            WashU Plasmodium EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library was constructed by R. Haywood. DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: L. David Sibley
            (sibley@borcim.wustl.edu), Washington University
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .407
                     /organism="Plasmodium falciparum 3D7"
                     /mol_type="mRNA"
                     /db_xref="taxon:36329"
                     /dev_stage="gametocyte (stage III-V)"
                     /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
                     /clone_lib="Plasmodium falciparum 3D7 gametocyte cDNA
                     library"
                     /note="Vector: pBluescript SK plus; Site_1: EcoRI; Site_2:
                     XhoI; The library was constructed by R Haywood. cDNAs were
                     synthesized  from ganetocyte poly(A)+ RNA by oligo d(T)
                     priming, size-selected  and directionally cloned into the
                     EcoRi (5' end) to XhoI (3' end)  sites of the Uni-ZAP XR
                     lambda vector (Stratagene). The primary  library was mass
                     excised as phagemid using the ExAssist helper  phage
                     (Stratagene). Clones were mass excised using the ExAssist
                     helper phage (Stratagene), the phagemids were precitptated
                     with  PEG 8000 and extracted with phenol/chloroform.
                     Phagemid DNA was  electroporated into DH10B cells. Clone
                     Availability: David Sibley,  Washington University."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 6;  Length 407;
  Best Local Similarity   85.0%;  Pred. No. 1.6e+03;
  Matches   17;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||| |||| |||||:|||
Db        162 AAATTTTTTAAAAGGGTCGA 181


RESULT 11
BI926889
LOCUS       BI926889                462 bp    mRNA    linear   EST 18-OCT-2001
DEFINITION  EST546778 tomato flower, buds 0-3 mm Lycopersicon esculentum cDNA
            clone cTOA31E1 5' end, mRNA sequence.
ACCESSION   BI926889
VERSION     BI926889.1  GI:16235836
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, 0-3 mm buds (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOA31E1"
                     /tissue_type="flower"
                     /dev_stage="0-3mm buds"
                     /clone_lib="tomato flower, buds 0-3 mm"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research; Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 4;  Length 462;
  Best Local Similarity   85.0%;  Pred. No. 1.6e+03;
  Matches   17;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||||| ||||||||||:| |
Db        356 AAATTGTTTATAAGGGTCTA 375


RESULT 12
BP527047/c
LOCUS       BP527047                494 bp    mRNA    linear   EST 28-SEP-2004
DEFINITION  BP527047 MAT001 Nicotiana tabacum cDNA clone BY11812, mRNA
            sequence.
ACCESSION   BP527047
VERSION     BP527047.1  GI:52830774
KEYWORDS    EST.
SOURCE      Nicotiana tabacum (common tobacco)
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   Matsuoka,K., Tashiro,G., Horiguchi,T., Demura,T. and Fukuda,H.
  TITLE     Profiling growth-phase dependent gene expression of tobacco BY-2
            cells by comprehensive microarray analysis
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ken Matsuoka
            Morphogenesis Research Group
            RIKEN Plant Science Center
            1-7-2 Suehirocho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9575
            Fax: 81-45-503-9573
```

```
            Email: by2@psc.riken.go.jp, URL:http://mrg.psc.riken.go.jp/strc/
            The cDNA library was constructed from mRNA isolated from lag (9 h),
            log (72 h) and stationary (7 days) old BY-2 cells.
            Seq primer: M13 forward.
FEATURES              Location/Qualifiers
     source           1. .494
                      /organism="Nicotiana tabacum"
                      /mol_type="mRNA"
                      /cultivar="Bright Yellow No.2"
                      /db_xref="taxon:4097"
                      /clone="BY11812"
                      /cell_line="BY-2"
                      /clone_lib="MAT001"
                      /note="Vector: pGEM-T easy; primer: M13 forward; mRNA
                      obtained from lag, log and stationary phase cells"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 5;  Length 494;
  Best Local Similarity   85.0%;  Pred. No. 1.6e+03;
  Matches   17;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              |||||||||| |||||:| |
Db        492 AAATTCTTTAGAAGGGTCAA 473


RESULT 13
CE435953
LOCUS       CE435953                508 bp    DNA     linear   GSS 27-SEP-2003
DEFINITION  tigr-gss-dog-17000335866781 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE435953
VERSION     CE435953.1  GI:36713767
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES              Location/Qualifiers
     source           1. .508
                      /organism="Canis familiaris"
                      /mol_type="genomic DNA"
                      /strain="Standard Poodle"
                      /db_xref="taxon:9615"
                      /clone_lib="Dog Library"
                      /note="Site_1: BstXI; Libraries were prepared from
                      peripheral blood"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 508;
  Best Local Similarity   85.0%;  Pred. No. 1.6e+03;
  Matches   17;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              |||||||||||||| : ||
Db         15 AAATTCTTTATAAGGTTGGA 34


RESULT 14
CR154100
LOCUS       CR154100                572 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHPP347b24, genomic survey sequence.
ACCESSION   CR154100
VERSION     CR154100.1  GI:49932945
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 572)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES              Location/Qualifiers
     source           1. .572
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10090"
                      /clone="MHPP347b24"
                      /clone_lib="MHPP"
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 572;
  Best Local Similarity   85.0%;  Pred. No. 1.7e+03;
  Matches   17;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAATTCTTTATAAGGGUCGA 20
              |||||||||||||| :| |
Db        209 AAATTCTTTATAAGGATCCA 228


RESULT 15
CD865385
LOCUS       CD865385                588 bp    mRNA    linear   EST 11-JUL-2003
DEFINITION  AZO2.073N10R000926 AZO2 Triticum aestivum cDNA clone AZO2073N10,
            mRNA sequence.
ACCESSION   CD865385
VERSION     CD865385.1  GI:32549201
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 588)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES              Location/Qualifiers
     source           1. .588
                      /organism="Triticum aestivum"
                      /mol_type="mRNA"
                      /cultivar="recital"
                      /db_xref="taxon:4565"
                      /clone="AZO2073N10"
                      /tissue_type="root"
                      /clone_lib="AZO2"
ORIGIN
```

Query Match 84.0%; Score 16.8; DB 6; Length 588;
Best Local Similarity 85.0%; Pred. No. 1.7e+03;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 AAATTCTTTATAAGGGUCGA 20
              ||| |||||||||||:| |
Db        157 AAAATCTTTATAAGGGTCTA 176
```

Search completed: March 17, 2005, 11:07:55
Job time : 1389.27 secs